Journal of the Association for Information Systems (2018) **19**(8)**,** 689-715
doi: 10.17705/1jais.00506

RESEARCH PAPER

ISSN 1536-9323

Don't Even Think About It! The Effects of Antineutralization, Informational, and Normative Communication on Information Security Compliance

Jordan B Barlow[1], Merrill Warkentin[2], Dustin Ormond[3], Alan R. Dennis[4]

[1]University of St. Thomas, jordan.barlow@stthomas.edu
[2]Mississippi State University, m.warkentin@msstate.edu
[3]Creighton University, dustinormond@creighton.edu
[4]Indiana University, ardennis@indiana.edu

Abstract

Organizations use security education, training, and awareness (SETA) programs to counter internal security threats and promote compliance with information security policies. Yet, employees often use neutralization techniques to rationalize noncompliant behavior. We investigated three theory-based communication approaches that can be incorporated into SETA programs to help increase compliance behavior: (1) informational communication designed to explain why policies are important; (2) normative communication designed to explain that other employees would not violate policies; and (3) antineutralization communication designed to inhibit rationalization. We conducted a repeated measures factorial design survey using a survey panel of full-time working adults provided by Qualtrics. Participants received a SETA communication with a combination of one to three persuasion statements (informational influence, normative influence statement, and/or an antineutralization), followed by a scenario description that asked for their intentions to comply with the security policy. We found that both informational (weakly) and antineutralization communication (strongly) decreased violation intentions, but that normative communication had no effect. In scenarios where neutralizations were explicitly suggested to participants, antineutralization communication was the only approach that worked. Our findings suggest that we need more research on SETA techniques that include antineutralization communication to understand how it influences behavior beyond informational and normative communication.

Keywords: Information Security, Neutralization, Training, Compliance, Normative Influence, Informational Influence, Factorial Survey

Anthony Vance was the accepting senior editor. This research article was submitted on January 15, 2016 and went through two revisions.

1 Introduction

As knowledge sharing and online transactions among individuals and organizations increase, information security increasingly becomes a strategic issue. Although organizations must protect against vulnerabilities from outside attacks (Ransbotham & Mitra, 2009) and comply with external security and privacy rules (Wall, Lowry, & Barlow, 2016), many security vulnerabilities arise from the actions of employees. Through the use of organizational sanctions and security education, training, and awareness (SETA) programs, security professionals actively battle security incidents (Jenkins & Durcikova, 2013) by encouraging employees to perform security-related behaviors, such as updating software, avoiding questionable emails, and using strong passwords (Anderson & Agarwal, 2010; Bulgurcu, Cavusoglu, & Benbasat, 2010; Johnston & Warkentin, 2010; Johnston, Warkentin, & Siponen,

2015; Karjalainen & Siponen, 2009; Liang & Xue, 2009; Liang & Xue, 2010) .

Modern organizations develop or purchase SETA programs designed to reinforce acceptable use guidelines and emphasize the potential consequences of information security policy violations (D'Arcy, Hovav, & Galletta, 2009). However, traditional SETA approaches are often ineffective in preventing violations (Siponen & Vance, 2010), so it is imperative that we explore other approaches to designing SETA programs and the way they communicate policies to better persuade employees to comply (Johnston, Warkentin, McBride, & Carter, 2016; Johnston et al., 2015). Many researchers conclude that managers should effectively communicate security-related concepts to their employees (Boss, Galletta, Lowry, Moody, & Polak, 2015; Siponen & Vance, 2010; Willison, Warkentin, & Johnston, 2018), yet little research empirically examines how such communication can affect later security behavior. Moreover, despite the extensive body of research on SETA, little research has examined the use of *periodic short communication* from management about the importance of complying with information security policies and actually applying what is taught in the formal SETA process.

Much like the successes from using so-called "nudges" by behavioral economists such as Tversky and Kahneman (1981) and Thaler and Sunstein (2008) for promoting tax compliance and healthy lifestyle decisions, such short periodic communications may serve as reminders regarding expected behaviors and may be critical to ensure greater security policy compliance. Therefore, in this paper, we focus on short SETA communications designed to augment SETA education and training programs. We do not address the content or effectiveness of these detailed programs. For an overview of SETA education and training approaches, see Puhakainen and Siponen (2010). Because annual SETA training effectiveness decays over time, some employers and software vendors have begun to implement real-world short communications. Providence Health and Services, a hospital chain on the U.S. west coast, has replaced office notes with ones that say "protect confidential information" and other reminders. The SANS Institute distributes post-it notes that include the reminder "do not write your password here." Commercial web browsers now utilize security warnings displayed to users who may surf to the wrong site (Anderson, Vance, Kirwan, Eargle, & Jenkins, 2016; Vance, Anderson, Kirwan, & Eargle, 2014). Many sites now provide instant feedback on the strength of newly formed passwords, which has been shown to have a positive impact on user security behavioral outcomes (Ciampa, 2013; Ur et al., 2012). Finally, studies have shown the value of general security pop-ups and warnings when users perform various actions (Akhawe & Felt, 2013).

Despite the widespread use of SETA programs that are designed to increase awareness of security policies and often emphasize the sanctions for violations, employees are often noncompliant (EY, 2017). Although some security incidents arise from deliberate employee actions, many vulnerabilities originate from nonmalicious security policy violations, such as sharing passwords. These nonmalicious violations may be due to human error, negligence, or poor training. The Global State of Information Security Survey 2015 indicated that employees remain the most cited perpetrators of security incidents. This survey of IT executives from 154 countries suggested that current employees and service providers/consultants were responsible for over 50% of reported incidents, with current employees being the most cited culprit at 35% (www.pwc.com/gsiss2015). Ernst and Young's Global Information Security Survey 2017 mirrored these findings: 55% of responding firms said careless or unaware employees were the greatest vulnerability they faced (EY, 2017).

When faced with a situation to comply or not comply with a security policy, employees assess potential consequences and often minimize perceptions of negative consequences by rationalizing that their actions are necessary to accomplish some higher objective (Sykes & Matza, 1957). For example, employees may believe that a security policy hinders their job performance, and because job performance is more important than security, it makes sense to violate the policy (Post & Kagan, 2007). More than a dozen such neutralization techniques have been identified (Willison & Warkentin, 2013) and many have been shown to have a greater impact on violation intentions than organizational sanctions; employees violate IT policies even when they know there are sanctions (Siponen & Vance, 2010). Through the use of short communication statements, we investigate three distinct approaches designed to increase compliance. The first two are drawn from prior research on persuasive communication and the third directly attacks neutralization.

Previous research has identified two main types of communication that can persuade someone to comply with a request for action: informational influence and normative influence (Burnstein & Vinokur, 1973). Information influence provides information and reasoned arguments in favor of compliance (e.g., by explaining why compliance is important) while normative influence provides information about others who are complying (e.g., by showing that

compliance is the organizational norm). Both informational and normative influence have been extensively researched in information systems, including topics such as system resistance (Kim & Kankanhalli, 2009), technology use and usefulness (Lewis, Agarwal, & Sambamurthy, 2003; Liang, Saraf, Hu, & Xue, 2007), viral marketing (Subramani & Rajagopalan, 2003), and computer-mediated communication (Sussman & Siegal, 2003). Prior information security studies have examined these as antecedents of information security behaviors. For example, Liang and Xue (2009) proposed that normative influence affects an individual's evaluation of IT threats and safeguarding measures as well as the motivation to avoid these threats; Bulgurcu et al. (2010) successfully demonstrated that both informational and normative influence have an impact on security awareness, attitudes, and intentions; Puhakainen and Siponen (2010) examined informational influence in training programs and determined it improved attitude towards and compliance with security policies, and Barlow et al. (2013) established the impact of informational influence on intentions to violate security policies. However, few studies have proposed and tested practical interventions that leverage these to influence compliance behavior (Lebek, Uffen, Breitner, Neumann, & Hohler, 2013).

Both information influence and normative influence focus on the situation and the desired behavior, not neutralization. Thus, a third possible source of influence directly targets the neutralizations that employees use to talk themselves out of complying, not the decision situation. This form of persuasive communication attacks neutralization techniques directly by communicating that neutralization is unacceptable, rather than using information influence or normative influence to focus on the situation or the policy. Previous studies support the argument for proper training against neutralization. For example, training that explains reasons for inequity is successful in preventing the use of neutralization techniques (Greenberg, 1990; Siponen & Vance, 2010). Siponen and Vance (2010) further suggest that organizations should help employees understand the consequences of not adhering to IS security policy and that rationalizations are unacceptable, but little research has examined this idea empirically. Such trainings could emphasize that there are inappropriate ways to respond to certain situations. Many information security policy studies have focused on what managers can do to effectively motivate compliance with information security policy; this study, however, focuses on how to effectively deter the individual from engaging in deviant behaviors resulting from rationalizations.

This leads to the following research questions: (1) *Can persuasive communication statements increase compliance with information security policies?* (2) *Which form(s) of persuasive communication have a greater impact on compliance with information security policies?* In this paper, we examine the effects of normative, information, and antineutralization SETA communication approaches in the context of rationalizing policy violations. Our results show that only two were effective: information influence and antineutralization communication. In situations where the probability of neutralization was high (i.e., the treatments in which we suggested neutralization arguments to the participants), only antineutralization communication was effective; neutralization overwhelmed information and normative influence.

The remainder of the paper begins by presenting the theoretical background that motivates our conceptual model and hypotheses. Next, we describe the factorial survey design and statistical analysis used to test these hypotheses. Finally, we discuss the results, implications, and limitations of this study and suggest potential future research opportunities.

2 Theoretical Background

When employees face a situation requiring them to choose to comply with or violate a security policy, they consider the information about the situation, the norms of others, and the potential reasons to rationalize away the need to comply (Barlow, Warkentin, Ormond, & Dennis, 2013). Thus, SETA programs that include communications using information influence, normative influence, and antineutralization approaches may increase security policy compliance. Two of these techniques are based on fundamental social influence theories: informational influence and normative influence (Burnstein & Vinokur, 1973), both of which have been shown as important internal motivators of security behavior (Bulgurcu et al., 2010; Liang & Xue, 2009; Puhakainen & Siponen, 2010). Informational influence theory argues that individual behavior is influenced by relevant information, such as the outcomes of the behavior, separate from any sanctions (Aronson, Wilson, & Akert, 2005; Burnstein & Vinokur, 1973; Shaw, 1981). For example, information influence could argue that sharing passwords increases the risk that someone could steal unauthorized information. Normative influence theory argues that individuals conform to norms of others to preserve a favorable self-presentation (Aronson et al., 2005; Burnkrant & Cousineau, 1975; Kaplan & Miller, 1987). For example, normative influence could argue that 99% of employees say they never will share a password. The third technique, which we call antineutralization

communication, directly addresses the temptation to neutralize, rather than the situation itself. For example, antineutralization statements could say that although some people believe that sharing passwords causes no harm, this is false; there are no circumstances where sharing passwords is justified.

Our study extends prior research by adapting these three approaches (informational influence, normative influence, and neutralization) into short SETA communication statements designed to reduce the intention to violate security policies such as sharing passwords. In other words, using theoretical levers informed by prior theory and empirical research related to security violation intentions, we designed SETA interventions to reinforce acceptable behaviors. Our focus is on situations in which it would be easy to rationalize security violations, so we also investigate the extent to which neutralization influences violation intentions. Figure 1 presents our conceptual model. The sections below describe each aspect in more detail.

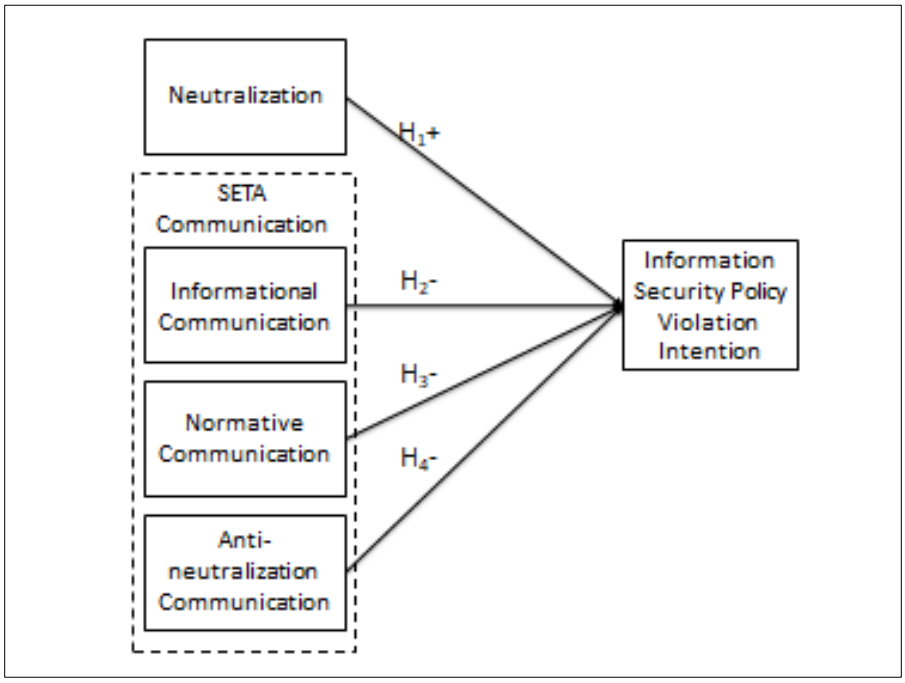

Figure 1. Conceptual Model and Hypotheses

2.1 Neutralizing Theory

Neutralization is defined as the use of rationalizations when violating a policy (Sykes & Matza, 1957). Neutralization theory, which originated from criminology research, identifies various forms of rationalization, termed "neutralization techniques" (Sykes & Matza, 1957). Individuals apply these techniques in noncriminal rule-breaking actions (Pershing, 2003). The original list of five neutralization techniques identified by Sykes and Matza (1957) has been extended to a list of 17 different neutralization techniques (Willison & Warkentin, 2013). Two examples are denial of injury and defense of necessity. In the instance of denial of injury, offenders perceive their behavior has no direct harmful consequences to the victim (Henry, 2009; Sykes & Matza, 1957); therefore, they rationalize their noncompliant behavior saying, "I didn't really hurt anybody." An individual rationalizing their behavior using the defense of necessity technique believes that he or she has no choice under the given circumstances but to engage in certain behavior (Cromwell & Thurman, 2003; Minor, 1981). For example, consider a security policy that discourages employees from writing down passwords, yet because they cannot remember them, they write them down. Alternatively, consider a policy that requires a data encryption procedure for all data transferred to a USB drive, but employees feel is too time-consuming to perform; employees could use neutralization techniques to justify violating security policies for "practical" reasons.[1]

Siponen and Vance (2010) applied neutralization theory to information security policy research; they

[1] We use USBs as an example, although we examined only password security in our experiment.

investigated specific neutralization techniques and found them to be more powerful than sanctions in predicting employee violations of information security policies. Subsequent research confirms that neutralization techniques are an important predictor of information security violations (Warkentin, Willison, & Johnston, 2011; Willison & Warkentin, 2010; Willison & Warkentin, 2013). Further, Willison and Warkentin (2013) call for more research into the role of neutralization in information security violations.

Neutralization theory proposes that when an individual considers violating a rule, cognitive dissonance exists between perceptions that an action is justifiable and knowledge that policy prohibits the action. In other words, the person cognitively reasons that the potential action is simultaneously wrong (according to the policy) and right (for some justifiable, context-specific reason) (Festinger, 1957). Neutralization techniques resolve, or at least reduce, the cognitive dissonance by changing the individual's perceptions of the negative consequences of violating the rule. By reducing perceptions of negative consequences, the individual chooses to violate the rule because there is less dissonance between the positive and negative outcomes of the behavior. Essentially, neutralization reduces or eliminates the perceived negative consequences associated with formal sanctions, informal sanctions, and/or shame from violating the policy (Paternoster & Simpson, 1996; Siponen & Vance, 2010). Neutralization may even cause violators to perceive their behavior to be less risky. Though research has already shown the power of neutralization in affecting security behavior intentions, we hypothesize and test this effect as a baseline for our remaining hypotheses. In essence, we first test that neutralization will have an effect in our specific study to validate the power of influence statements in situations where individuals rationalize insecure behavior. Therefore, we hypothesize that:

H1: Employees using neutralization are more likely to form intentions to violate an information security policy.

2.2 Persuasive Information Security Communication

Given the considerable impact that neutralization has on violation intentions, the purpose of this study was to understand which types of influence communication induce good security behaviors while reducing the effects of neutralization. A vast amount of information security research has determined that deterrent-focused communication reduces violation intentions. Barlow et al. (2013) demonstrated that neutralization-focused communication is just as effective as deterrent-focused communication. Our study extends this research by separately examining three different types of communication (informational, normative, and antineutralization) and testing their individual effects on intentions to violate information security policies. Both informational and normative statements are adapted from prior theory whereas antineutralization statements are a third theory-based approach we created for this study.

2.2.1 Informational Communication

Individuals can be persuaded by detailed information about alternatives when they deeply consider new information about them (Aronson et al., 2005; Dennis, 1996; Dennis, Hilmer, & Taylor, 1998). Informational influence occurs when "behavior is based on a personal evaluation of the information provided" (Subramani & Rajagopalan, 2003). Informational influence is driven by the direct evaluation of the costs and benefits of the behavior (Subramani & Rajagopalan, 2003). With informational influence, individuals actively consider information given to them and incorporate the information into their existing mental schemas (Petty & Cacioppo, 1986). When the goal is to influence individuals to abstain from engaging in a specific behavior, then information that highlights the costs of that behavior may influence them to abstain. Within the IS security literature, some studies have examined the effect of perceptions of benefits and costs of noncompliance and found that such perceptions make a difference (Bulgurcu et al., 2010; Xue, Liang, & Wu, 2011). Thus there is a need to understand if such perceptions can be intentionally influenced by informational communication to the employee.

Our focus is on the direct costs of the security violation behavior itself, not on indirect costs, such as sanctions designed to deter such behavior. We focus on direct costs because other research has demonstrated sanctions to have some, albeit limited, effect and because sanctions are theoretically distinct from direct consequences of the behavior; that is, they are externally imposed, rather than being directly inherent to the behavior itself (Johnston et al., 2015; Son, 2011). Thus rather than simply stating a policy and expecting employees to comply with it, SETA communications following this theoretical path include statements that give information about the consequences of engaging in a proscribed behavior or failing to engage in a required behavior. For example, communication regarding a password policy prohibiting password sharing would explain how unauthorized access to private data could occur through sharing passwords. Similarly, communication about a policy requiring USB drives to be encrypted would explain how unauthorized access to private data could occur if the USB were not encrypted and subsequently lost. The goal of these communications is to prompt the individual to engage in a thoughtful

consideration of the information behind the policy so that he or she concludes that the behavior desired by the policy is the appropriate behavior. Ideally, this conclusion would be strong enough for the individual to perform the desired behavior even in the absence of a policy.

Empirical evidence suggests that individuals assess information supporting a security policy when they consider engaging in a policy-violating action (Puhakainen & Siponen, 2010) and use moral reasoning to come to a decision (Myyry, Siponen, Pahnila, Vartiainen, & Vance, 2009). Several studies providing information using fear appeals or dialogue and reflection show that providing information about a threat and a response that can mitigate the threat influences compliance behavior (Albrechtsen & Hovden, 2010; Johnston & Warkentin, 2010; Puhakainen & Siponen, 2010). Such studies suggest that by providing appropriate supporting information about the reasons for security policies, organizations may positively influence employee security behaviors (Liang & Xue, 2009). Although these prior studies demonstrate that information influences security behavior, they have not examined this influence in situations where the potential benefits of security policy violations are so great that the decision calculus is altered and employees are more likely to neutralize their violation behavior.

A search of prior research reveals only one study that actually examined whether providing information about the consequences of security policy violations influences violation behavior. In an action research study of a Finnish company, researchers started by delivering a three-session multihour training program on the policy of encrypting emails containing confidential information (Puhakainen & Siponen, 2010). The first training session was a presentation of the risks related to the use of email, followed by employees searching their own emails to find sample emails that contained confidential information as defined by the security policy, and then asking employees to identify the consequences if this information were to be discovered by competitors. The second session was only for nontechnical users and focused on how to use the encryption software to email encrypted documents, including sharing encryption passwords with the receiver over the phone. The third session reviewed the issues covered in the first two sessions.

By the end of the sessions, nine employees understood the information describing the rationale for the policy and expressed a willingness to comply with it, whereas five other employees felt the policy "was still too far removed from" the company's other activities and thus were not motivated to comply. A broader second round of information-based security training was delivered to the entire company once a month. After three months, this program was determined to be *ineffective* in influencing employees' motivation to comply with security policy. At that point, the CEO stepped in and made security his priority in monthly meetings. Two months later, five employees reported that they were more motivated to comply.

One important difference between the Puhakainen and Siponen (2010) study and ours is the nature of the SETA communication. Puhakainen and Siponen (2010) examined the effects of a long face-to-face training session. In contrast, we are interested in a specific type of short written reinforcing communication, sometimes characterized as "nudges" (Thaler & Sunstein, 2008). Though there are some similarities, brief SETA communication is fundamentally different from in-depth SETA training, making it unclear about the extent to which the results from Puhakainen and Siponen (2010) would generalize to our setting.

Further, the Puhakainen and Siponen (2010) study did not analyze whether the informational persuasion technique would be strong enough in situations where employees might commonly neutralize their behavior. Neutralization theory posits that individuals use rationalizations when faced with conflicting values (i.e., benefits of violating vs. risks of violating). Informational communication should help resolve the cognitive dissonance individuals feel when they are in a situation where noncompliance would seemingly result in some benefit (e.g., convenience), by reminding them of the reason for the policy and increasing the perception of compliance benefits.

In summary, theory suggests that communicating information about the consequences of security policy violation prior to a violation decision should influence behavior, but the empirical evidence (which studied training, not communication) is unclear. We rely on the former argument and hypothesize:

H2: Employees receiving written communication containing informational influence statements are less likely to form intentions to violate an information security policy.

2.3 Normative Communication

Normative influence occurs when individuals use the beliefs and actions of others as a guide for their own behavior, rather than using their own understanding of the available information to form their decision (Aronson et al., 2005). Normative influence is based on the expectation of a need to conform to the others' behavior (Subramani & Rajagopalan, 2003). Individuals who experience normative influence comply due to the desire for social support or to avoid social sigma (Ajzen, 2002; Kim & Kankanhalli, 2009). Given the social expectations involved with

normative influence, behavior induced by normative influence is more likely to be discontinued when it is no longer observable (Subramani & Rajagopalan, 2003).

Normative influence has been theorized to influence decisions about whether or not to violate security policy (Liang & Xue, 2009). Past empirical research suggests that normative beliefs (e.g., subjective norms) play a significant role in the decision to violate a policy (Herath & Rao, 2009; Ifinedo, 2012; Ifinedo, 2014; Siponen, Mahmood, & Pahnila, 2014), a role as important as informational influence (Bulgurcu et al. 2010). Several studies (Chatterjee, Sarker, & Valacich, 2015; Lowry & Moody, 2014; Siponen et al., 2014) examined individuals' self-reported post hoc perceptions of the factors influencing their behavior after making a decision about whether to not to violate a policy. That is, when asked after the fact, some users say normative influence affected their security decisions. These studies do not directly address the question of whether providing normative information before a decision changes intentions or behavior, but they suggest that intentionally crafting a message designed to provide normative influence before the decision would influence behavior.

We argue that providing normative information that purports to describe what others would do in a similar situation will influence security violation decisions. For example, individuals who are motivated by normative information are more likely to comply with information security policy when subjective norms suggest that many other employees would be likely to comply (Herath & Rao, 2009; Ifinedo, 2014). Normative influence is a factor driven by an individual's understanding of his or her own reference group, such as peers (Bulgurcu et al., 2010). It is unclear whether SETA communication containing statements about others would influence behavior. Nonetheless, normative beliefs are often very powerful, and normative influence from SETA communication may help employees resolve the cognitive dissonance felt in situations where there are benefits for both compliance and noncompliance. Thus, we posit that:

H3: Employees receiving written communication containing normative influence statements are less likely to form intentions to violate an information security policy.

2.3.1 Antineutralization Communication

With information influence and normative influence, the facts about the situation (information influence) and perceptions of what others would do (normative influence) drive compliance decisions. Neutralization involves rationalizing away the information influence and normative influence factors that support compliance so that they become less important than the factors favoring noncompliance. Neutralization is often more powerful than information and norms, and even more powerful than deterrent sanctions (Silic, Barlow, & Back, 2017; Siponen & Vance, 2010). Employees often use neutralization when they encounter strong sanctions, especially when they feel their organization is treating them unjustly (Warkentin et al., 2011; Willison & Warkentin, 2013).

In this section, we focus on a third potential type of security communication that focuses on the neutralization process, not the underlying situation in which employees find themselves. Antineutralization communication does not address the situation. Rather it is separate and distinct in that it focuses on neutralization behaviors that are not tied to the specific situation.

Neutralization behaviors are often driven by cognitive dissonance. Good security behaviors often entail a cost to the employee by requiring extra effort (e.g., logging off a computer when they leave their desk, encrypting a USB drive, or choosing new complex passwords that are hard to remember for each resource). When employees perceive extra costs from complying with a policy, they may experience cognitive dissonance between those factors that motivate compliance and noncompliance. This confounds the compliance decision calculus, causing them to be tempted to violate the policies. Employees may consider rationalizing their behavior, which has been shown to strengthen the relationship between perceived injustice and computer abuse intentions (Warkentin et al., 2011; Willison et al., 2018) as well as directly influence deviant behavior such as intention to violate policies (Siponen & Vance, 2010) and cyberloafing (Lim, 2002).

Justifications for negative security-related behavior may be based on certain heuristics or biases, such as anchoring, optimism bias, loss aversion, etc. (Tsohou, Karyda, & Kokolakis, 2015). For example, with the denial of injury neutralization technique, employees may be biased in their judgments of the harm that could be caused by potential security violations. Training against such processes should allow employees to more carefully think through their decision-making process by making them aware of their biases. SETA programs have been shown to effectively deter misuse, and the IS security literature calls for more research to develop practical SETA techniques (D'Arcy et al., 2009; Dhillon, 1999; Posey, Roberts, & Lowry, 2015; Straub & Welke, 1998). Integrating training against neutralizations into these SETA programs and other training materials would help users become more cognizant of thought processes that may have negative consequences on both the organization and the individual; therefore,

they may be effective in deterring individuals from acting on these thoughts.

Because employee rationalizations make violation decisions more attractive by serving to eliminate such cognitive dissonance, organizational training against neutralization can be used to discourage these justifications (Barlow et al., 2013; Willison et al., 2018). Antineutralization communication focuses on this cognitive dissonance by arguing that employees should never rationalize such violations. Antineutralization communication argues that there are no situational ethics; that is, there are no situations in which violations can be justified. Note that antineutralization communication is not specific to the situation because its focus is neutralization (in contrast, both information influence and normative influence focus on the specific situation).

Antineutralization communication can be neutralization specific (e.g., targeting defense of necessity by arguing there is always an alternative to violating a policy or targeting denial of injury by arguing there is always the possibility for harm). By explicitly recognizing the process of rationalization, antineutralization communication aims to encourage employees not to rationalize to reduce dissonance. Rather, employees will be persuaded that when rationalizations come to mind, the correct course of action is to ignore them and comply with the policy.

Antineutralization statements will have the greatest effect—perhaps the only effect—when neutralization is strong by acting to reduce its effects on intentions; without neutralization, antineutralization communication is likely to have little effect, although there is the opportunity for neutralization in most compliance situations because compliance requires additional effort. Thus, communication focused on mitigating neutralization may reduce rationalization behaviors and ultimately intentions to violate security policy.

One previous study (Barlow et al., 2013) tested the effects of "neutralization focused" communication. The study used a factorial survey of 360 respondents to examine the effects of antineutralization communication, deterrence communication, and message framing (positive vs. negative). They found that antineutralization communication significantly reduced violation intentions and was comparable in impact to deterrence communication. The antineutralization treatment also included information-based statements, so it was not a test of antineutralization communication separate from information-based communication; rather, it was a test of the combined effects of both information and antineutralization. Therefore, it is still unclear whether antineutralization communication alone can cognitively influence individuals' security intentions in a different manner than the two more traditionally studied forms of influence (informational and normative). Therefore, we hypothesize that:

H4: Employees receiving written communication containing antineutralization statements are less likely to form intentions to violate an information security policy.

3 Method

We applied the factorial survey method design to test our hypotheses (Jasso, 2006; Rossi & Anderson, 1982; Shlay, Tran, Weinraub, & Harmon, 2005; Wallander, 2009). In the scenario-based factorial survey design, participants read several unique scenarios that contain a subset of the experimental treatments and then answer survey questions based on their perceptions of each scenario. Various versions of the baseline scenario embed language that orthogonally represents the independent variables under investigation (thereby eliminating the possibility of multicollinearity between predictor variables), and the respondent is asked if he or she would act in the same way as the scenario character (thus, intentions serve as the dependent variable). Security and business ethics researchers often use scenario-based methods (Herzog, 2003; Seron, Pereira, & Kovath, 2006; Trevino, 1992; Weber, 1992) because it is difficult, if not impossible, to measure actual deviant behavior in the workplace by observation or direct questioning, in part due to social desirability bias. Instead, participants report whether they would act in a similar manner as a character in the scenario (Harrington, 1996; Trevino, 1992), thereby removing feelings of incrimination for violating behaviors while still capturing intentions (Crossler et al., 2013). Vance, et al. (2015, p. 353) point out that scenarios "afford an indirect way of measuring the intention to commit unethical behavior" by using hypothetical terms. The scenario technique is the most common method in studies of ethical issues (O'Fallon & Butterfield, 2005) and is increasingly applied to study IT security policy violations (Argelaguet, Kulik, Kunert, Andujar, & Froehlich, 2011; Barlow et al., 2013; Goel, Williams, & Dincelli, 2017; Guo, Yuan, Archer, & Connelly, 2011; Johnston et al., 2016; Trinkle, Crossler, & Warkentin, 2014; Willison et al., 2018). However, some of these security studies used scenarios without the use of the factorial survey methodology. The factorial survey method utilizes a full factorial of all realistic combinations of all levels and dimensions of each variable being investigated, whereas scenario-based surveys may not incorporate the complete calculus of the relationships between all model components. Furthermore, multiple versions of the scenario or vignette are repeated within the study, possibly evaluated numerous times by numerous respondents, which strengthens the value of the

factorial survey method. For further elaboration of this method, see Vance, Lowry, & Eggett (2015) and Willison et al. (2018).

Before distributing the scenario-based survey, we convened an expert review panel, as suggested in previous research (Straub, Boudreau, & Gefen, 2004), to ensure realism, content validity, and face validity. The panel consisted of experts in instrumentation and scale development, as well as experienced security experts providing feedback on generalizability and realism of scenarios. As a result of the panel review, we revised the survey instrument and scenarios to be more comprehensive and realistic; our revisions also reduced ambiguity and potential survey fatigue (Lanza, 1988; Lauder, 2002). We then conducted a small pilot study with a convenience sample to confirm discriminant and convergent validity before primary data collection commenced. Survey metadata confirmed completion time estimates and other factors.

3.1 Participants

Participants were recruited through Qualtrics, a survey and Internet panel provider firm. Qualtrics recruited U.S. participants who passed filter questions indicating they had full-time experience in a workplace using computers and security policies. All subjects completed the survey anonymously. To prevent survey fatigue and reduce learning effects and hypothesis guessing, each participant completed a random set of two (out of 24 possible) scenarios. After rigorous manipulation checks and quality checks, our final data sample consisted of 200 participants, or 400 scenario responses. See Appendix C for more details on recruitment and participant filtering procedures.

3.2 Task

Each participant completed the online survey at his or her convenience. The respondents read two scenarios that detailed violations of information security policies in familiar situations. First, the scenario introduced a hypothetical company that had established information security policies and procedures related to password sharing. Next, respondents read a (randomly assigned) combination of up to three persuasion statements—a normative influence statement, an antineutralization statement, and an informational influence statement. Then, the scenario described a situation where a particular employee of the hypothetical company violated the information security policy by sharing his computer password because the fictional individual concluded that there were benefits to noncompliance. Finally, we collected responses to various items, including the violation intention of the participant if they experienced similar circumstances. Each participant repeated this task with a second (randomly generated) scenario. See Appendix A for the full text of the scenario combinations.

3.3 Experimental Treatments

To test our hypotheses, we conducted a 2 (normative influence statement present or not) x 2 (antineutralization statement present or not) x 2 (informational influence statement present or not) x 3 ("denial of injury" neutralization statement, "defense of necessity" neutralization statement, or no neutralization statement) factorial design.

In the normative influence treatment, the scenario reported that a large majority of employees would comply with the policy in all situations (as determined by a company survey). In the antineutralization treatment, the scenario stated that employees should not violate the policy even when they perceive justification of the action as an option. In the informational influence treatment, the scenario stated that the reason for the policy was that sharing passwords may result in serious consequences, such as malicious deviant behavior by the employee with whom the password is shared. Finally, in the neutralization statement treatment, a sentence at the end included the rationalization for violating an IT policy. Some versions of this treatment used the "denial of injury" technique, where the employee feels that no harm would result from violation (Siponen & Vance, 2010; Warkentin et al., 2011). Other versions used the "defense of necessity" technique, where the employee feels that violation is necessary for a greater cause (Siponen & Vance, 2010). We chose these techniques because they are particularly relevant to password sharing; our expert panel reviewers agreed that these rationalizations were realistic in the scenarios. We note that not all neutralization techniques are equal; some are more powerful than others depending on the individual and the situation (Barlow et al., 2013).

3.4 Dependent Variable Measurement

We asked each respondent to rate the likelihood that he or she would violate the given security policy under similar circumstances (Paternoster & Simpson, 1996; Siponen & Vance, 2010). To avoid reliability issues, we used four items for the dependent variable (see Appendix B), with each item using a fully anchored five-point Likert-type scale ranging from strongly disagree to strongly agree. The Cronbach's alpha was .929, indicating adequate reliability.

3.5 Experimental Procedures

Participants received a link to participate in the study from Qualtrics. After completing a consent statement, each person answered two filter questions gauging

whether he or she had experience in a company with computers and formal policies (see Appendix B). If the participants answered negatively to either of the questions, the survey ended.

Following the filter questions, participants viewed a scenario, randomly assigned from the full set of scenarios, using the Qualtrics randomization feature, immediately followed by a set of manipulation check questions. We designed the manipulation check questions to ensure that the participant understood the scenario correctly in accordance with the experimental manipulations. Failure of manipulation check questions in experimental or survey research is highly correlated with lack of motivation and/or disregard to instructions from the researcher (Oppenheimer, Meyvis, & Davidenko, 2009). Data from participants who fail manipulation checks should thus be excluded from analysis (Oppenheimer et al., 2009). Participants received up to four manipulation questions, one each for the four dimensions of the study (i.e., antineutralization statement, normative statement, informational statement, and neutralization type). These manipulation questions are included in Appendix B. The participants only viewed manipulation check questions for the manipulations present in the scenario they experienced, so that the question would not induce bias or prime the participant.

After the manipulation check questions, participants responded to four dependent variable items and two additional items. First, a response set item (e.g., "select disagree as the response to this question") was used to ensure responses were based on attentive reading of the question rather than simply answering in patterns and not paying attention (Andrich, 1978; Kerlinger, 1973). Second, a realism item (i.e., "I could imagine a similar scenario taking place at my company") was used as a control variable (Siponen & Vance, 2010). After completing two scenarios, participants answered a set of demographic questions. (See Appendix B for all measures.) In the factorial survey method, participants often view and respond to many different scenarios. However, we decided to limit the number of scenarios to two per participant to reduce survey fatigue and learning effects.

3.6 Addressing Potential Bias

Because various forms of bias, including common method bias, are a serious concern for field studies and surveys (Podsakoff, MacKenzie, Lee, & Podsakoff, 2003), we followed the recommendations of Podsakoff et al. (2003) to address a number of specific potential bias threats, and to ameliorate their potential. A common issue that threatens field studies, especially in the information security field, is social desirability, the tendency to respond to questions in a culturally acceptable way (Podsakoff et al., 2003). We addressed social desirability by, first, using the scenario technique rather than direct questions. This technique increases the likelihood that participants will give true responses about information security policy violation intentions (Trevino, 1992). Second, participant responses were completely anonymous, as we did not receive any personally identifiable information from Qualtrics.

To address common method bias, we randomized the set of scenarios that each participant would receive to reduce tendencies to answer questions about each scenario based on previous scenarios. Another way we addressed common method bias was to use the response set question (Andrich, 1978; Kerlinger, 1973; Rennie, 1982) to ensure that participants would not simply give patterned, automatic responses. Further, we used manipulation check questions to ensure attention to response—any incorrect response to manipulation check questions resulted in discarding the responses for that scenario.

Another important issue was to ensure realism of the scenarios to encourage responses that are more valid. First, two expert panels reviewed the scenarios and questions to ensure realism. Second, we included the realism question in the survey to control for the effects of scenario realism (Siponen & Vance, 2010). Finally, the response set and realism questions were intermixed with dependent variable items to avoid grouping constructs (Podsakoff et al., 2003).

4 Results

The final data set consisted of 200 individuals, answering two scenarios each, resulting in a sample size of 400. Demographic information is summarized in Table 1. As a robustness check, we also repeated the analysis on data including participants who were filtered out of the final data set. The results of this data analysis are summarized in Appendix C.

Ordinary least squares regression (OLS) is a preferred technique for the factorial survey design (Rossi & Anderson, 1982; Shlay et al., 2005) due to the ease of interpreting coefficients. However, OLS requires a normal distribution of the dependent variable. In behavioral security research, the dependent variable often displays a skewed distribution because of the sensitive nature of admitting guilt to violating rules

Table 1. Demographic Information

Gender		**Work experience in years (mean = 24.2)**	
Female	121 (60.5%)	1-10	32 (16.0%)
Male	78 (43.3%)	11-20	48 (24.0%)
Unspecified	1 (0.5%)	21-30	58 (29.0%)
Age (mean = 45.4)		31-40	42 (21.0%)
20-29	29 (14.5%)	41-51	19 (9.5%)
30-39	40 (20.0%)	Unspecified	1 (0.5%)
40-49	47 (23.5%)	**Level of education completed**	
50-59	55 (27.5%)	High school	58 (29.0%)
60-71	29 (14.5%)	Undergraduate degree	83 (41.5%)
		Graduate degree	59 (29.5%)

The Kolgomorov-Smirnov (0.180; df = 400; p < 0.001) and Shapiro Wilk (0.870; df = 400; p < 0.001) tests of normality both indicated that our data are not normally distributed. Another assumption of OLS regression is independence of errors. Repeated-measures designs violate this assumption because responses from the same subject are likely to be correlated. As an alternative to OLS regression, Rossi and Anderson (1982) note that any multivariate technique that fits the data can be used for the factorial survey design. We chose to analyze the data using the "generalized linear mixed models" function in SPSS 24.0.0.0 (see syntax in Appendix F). This analysis approach allows for robust estimation to handle violations of distribution assumptions; it also allows for correlated observations (such as when participants view multiple scenarios) (Vance et al., 2015).

Table 2 lists the full results of the data analysis as well as a follow-up analysis with selected scenarios; we focus on the full results first. In our model, the base level can be interpreted as the scenarios where the manipulation statement was not present. For example, the parameters listed for the normative influence statement variable in the "Full results" columns of Table 2 indicate the difference between scenarios with this manipulation and those scenarios without it.

Table 2. Results for Intention to Violate the Security Policy

	Full results		**Only scenarios with neutralization present**	
	Estimate	p	Estimate	P
Intercept	-0.443	0.216	-0.026	0.949
Use of neutralization techniques	0.227*	0.026	n/a	n/a
Informational influence statement	-0.175†	0.092	-0.202	0.107
Normative influence statement	0.065	0.493	0.061	0.617
Antineutralization influence statement	-0.320**	0.001	0.369**	0.002
Order	-0.142*	0.010	-0.123	0.168
Realism	0.058	0.267	-0.013	0.835
Gender (Female)	0.092	0.465	0.085	0.572
Age	0.000	0.982	-0.005	0.676
Work experience	-0.005	0.624	-0.004	0.701
Education	0.097	0.558	0.077	0.698
Note: †p < 0.10; *p < 0.05; **p <0.01; *** p < 0.01;	n = 400 scenarios		n = 266 scenarios	

These results show that participants who viewed scenarios in which a neutralization technique was present were significantly more likely to form information security policy violation intentions (p = .026), providing support for Hypothesis 1. Participants who received a scenario containing informational statements were less likely to form information security policy violation intentions (p = .092), as were those receiving antineutralization statements (p = .001). However, there were no significant effects for normative influence statements (p = .493). These results provide support for

Hypotheses 2 and 4, but not Hypothesis 3. The raw data for intention scores are shown in Appendix D.

The two rightmost columns in Table 2 show the results when analyzing only the scenarios where a neutralization technique was used. That is, in addition to testing the effects of the various persuasion techniques on violation intentions in general, we also tested the effects of the techniques in situations where participants were highly prone to neutralize their behavior. In this situation, we see that antineutralization statements had a significant impact on behavior ($p = .002$), but information influence ($p = .107$) and normative influence statements ($p = .617$) did not. These results require some qualification to our conclusions about Hypothesis 2 that information influence statements affect behavior, because it is not supported when neutralization is high (and the significance levels were near 0.10 in the full data set). When neutralization is high, only antineutralization statements have an impact. Table 3 summarizes the support for each hypothesis.

Table 3. Summary of Hypotheses

H1. Use of neutralization techniques → Higher intentions to violate	Supported
H2. Informational influence statements → Lower intentions to violate	Weakly supported
H3. Normative influence statements → Lower intentions to violate	Not supported
H4. Antineutralization statements → Lower intentions to violate	Supported

To test order effects between the two scenarios that participants viewed, we used a dummy variable coded as whether the scenario was the first or second the participant had experienced. Scenario order had a statistically significant effect on intentions overall (but not for scenarios with neutralizations present), with first scenarios being rated with higher information security policy violation intentions than subsequent scenarios. This order effect indicates that as participants viewed additional influential communication in the scenarios, their information security policy violation intentions decreased. This indicates that viewing a second repetition of the security policy (without viewing possible neutralization statements) increased the strength of the security policy, while viewing both the policy and neutralization statements twice had no effect (i.e., they cancelled each other out); thus we conclude that the policy statement had the desired effect in increasing compliance (although that is not part of our research question). We also tested order as a moderator of the relationships between the main variables and the dependent variable; none of these tests was statistically significant, indicating that the order in which a person viewed the scenarios did not have an effect on other relationships in the model.

As a final check for order effects, we ran a separate model with only those scenarios first seen by a participant; this resulted in similar parameter estimates to the original model. However, the statistical significance was not as strong. Interestingly, the explicit use of neutralization techniques in the scenario was no longer statistically significantly related to intentions to violate the policy. Informational statements, which were significant at the $p<0.1$ level in the main results, did not have a statistically significant effect when analyzing only the first scenarios. However, it should be noted that the coefficient ($\beta = -0.188$) is actually greater than in the main results ($\beta = -0.175$), but cutting the sample size in half greatly increased the p value. That is, the effect size is similar, but the statistical significance is weak. See Appendix C.

The perceived realism of the scenarios did not have a statistically significant effect on how the participants rated their intentions. Furthermore, running the model on a data set that discards responses with low rated realism shows results with parameters and significance tests similar to the main model.

Because we used two types of neutralization techniques in the scenario for more generalizability, we conducted post hoc testing to investigate any differences between the effects of these techniques. Interestingly, we found that when testing them separately, the defense of necessity technique was significantly related to information security policy violation intentions ($\beta = 0.320$; $p = 0.008$), whereas the denial of injury technique was not ($\beta = 0.075$; $p = 0.534$). Participants were more likely to express intentions to violate the policy when they felt it was necessary than they were when they believed no one would be injured. However, the antineutralization was just as powerful in the denial of injury scenarios as in the defense of necessity scenarios, suggesting that whether we indicated a specific type of neutralization to the participant explicitly or not, the antineutralization statement influenced them to be less likely to violate. These results are summarized in Table 4.

Table 4. Results Comparing Denial of Injury and Defense of Necessity Separately to Baseline Scenarios with No Neutralization Statement

	Denial of injury (vs. baseline)		Defense of necessity (vs. baseline)	
	Estimate	p	Estimate	P
Intercept	-0.580	0.135	-0.762	0.080
Use of neutralization techniques	0.075	0.534	0.320**	0.008
Informational influence statement	-0.149	0.219	-0.168	0.187
Normative influence statement	0.123	0.321	-0.017	0.873
Antineutralization influence statement	-0.280*	0.021	-0.274*	0.013
Order	0.084	0.318	-0.380**	<0.001
Realism	0.112†	0.060	0.089	0.145
Gender (Female)	0.174	0.220	0.020	0.891
Age	-0.002	0.850	0.008	0.476
Work experience	-0.004	0.718	-0.009	0.440
Education	0.162	0.360	0.056	0.761
Note : †p < 0.10; *p < 0.05; **p <0.01; *** p < 0.01;	n = 259 scenarios		n = 275 scenarios	

We also analyzed all interaction effects between neutralization types and the hypothesized effects of informational, normative, and antineutralization statements and found no significant results (see Appendix E). In other words, the effects of the different types of persuasion were similar regardless of the specific neutralization type used in the various scenarios. It is possible that even in situations where we did not present an explicit neutralization statement in a scenario, participants came up with possible rationalizations of their own. Even when we presented an explicit neutralization statement, participants may have thought of other potential justifications as well. This result further strengthens the contribution of our findings in that we can be more confident that antineutralization statements are influential against multiple types of neutralization techniques. It would be difficult for organizations to use different techniques against different neutralizations that different employees might choose to use.

5 Discussion

5.1 Interpretation of Results

This study shows that the way organizations communicate security policies can increase compliance, over and above sanctions that are in place to deter violations, particularly in those situations where individuals consider benefits of noncompliance. First, above and beyond the enforcing behaviors implemented by organizations (e.g., sanctions, formal SETA programs), we show that *reinforcing* the need for secure behavior through short communications, including even brief informational statements that highlight the reasons why information security policies exist (i.e., the direct negative consequences of noncompliant behavior), increases the likelihood that employees will comply with the policy. When cognitive dissonance between violating information security policy and complying with information security policy exists, informational statements aid in decision making by reminding employees of reasons to comply. Such reminders should help employees decide on compliance based on a rational decision-making rather than justifying violation behavior through neutralization techniques or rash decisions. However, this effect is diminished in situations where an individual is more strongly tempted to neutralize. In scenarios where a neutralization technique was explicitly stated, the effect of informational communication was not statistically significant.

Second, statements specifically designed to counter neutralization techniques (i.e., "antineutralization") significantly reduced the intention to violate the policy. Antineutralization statements had the same effect whether informational statements were present or not, indicating that they were processed in a different manner than informational statements. Likewise, antineutralization statements had the same effect whether neutralization statements were present or not. Reinforcing communication that states that neutralization is unacceptable effectively combats rationalizations that lead to deviant behavior, whether rationalizations are explicitly triggered by the communication or spontaneously invoked by the users.

Interestingly, our results show that short normative statements in reinforcing communication do not reduce intentions to violate information security policy. For employees who contemplate rationalizing

violation, reading normative information did not increase compliance behavior. This finding is in contrast to previous research that has shown that internally held normative beliefs influence behavior (Bulgurcu et al., 2010); however, in situations where rationalization is possible and even probable, the effects of normative communication are not powerful enough to convince employees to comply with security policies. Normative influence may be powerful if it is a deeply held individual belief, but it is not as powerful in the form of short reinforcement statements from the organization. Normative influence tends to be more powerful when individuals are closer in "proximity" (whether space or time) to others whose norms they are considering (Latané, 1981). That is, norms may influence security behavior, but the normative influence individuals experience directly from their peers is likely stronger that reading normative information in SETA communication from the organization. While previous research has shown that normative influence plays a role in security behavior, our findings highlight the fact that knowing antecedents of behavior cannot directly translate into knowing how to implement meaningful interventions in practice.

5.2 Limitations

One limitation is that we studied short, focused SETA communications. We followed prior theory (and most empirical research) on information influence and normative influence, which has focused on how simple short communication can influence behavior. Though our use of theory here is a good match with the underlying theory and prior research, simple short communication is likely to have a weaker effect than the use of active learning techniques used in SETA training that are likely to stimulate deeper consideration of the message content. For example, compare our intervention, which took minutes, to the intervention of Puhakainen and Siponen (2010), which took hours.

Another potential limitation is that we measured compliance intentions directly after delivering information security communication to the participants. In reality, a longer period may pass between security communication and the decision to violate a policy. Future research could include distractor tasks between the SETA communication and the scenario assessing their behavioral impacts, or could assess behavior days or weeks after the SETA communication. Previous research has demonstrated that brief experimental interventions in normative influence communication can have an effect for at least a month, though the effect does attenuate over time (Zitek & Hebl, 2007), and that information influence communication can last longer than normative influence communication (Kaplan & Miller, 1987). Therefore, it may be possible to use research designs that have longer time spans separating the SETA communication and the measurement of intentions.

Further, the short messages were presented in the same order within the scenarios (1) because of limitations in the setup of the randomization of messages within the Qualtrics survey, and (2) because each message was placed within the scenario in a manner that made the wording of the scenario flow more effectively. Because the messages are short and appear close together, and because we included rigorous manipulation checks for each of the messages, we are confident that this would have little to no effect on the participants. However, there is always the possibility of a small effect.

An additional limitation may be introduced by testing the theories only within one specific security policy violation behavior, namely password sharing (and with only two potential neutralization strategies). Though we hope that our antineutralization statement was general enough to apply to many different types of neutralization techniques, our study only specifically examined two. Future research should examine this technique against other neutralization types because neutralization types work differently depending on the situation (Silic et al., 2017). Several studies have demonstrated the consistency of research results across multiple security scenarios. Johnston et al. (2015) found the same results from studying user reactions to password theft, to USB memory card loss, and from data theft from not logging off or from locking workstations. Siponen, Vance, & Willison (2012) similarly demonstrated consistent results across four unique security scenarios, and Siponen and Vance (2010) revealed somewhat consistent results from the three situations—USB drive loss, workstation logout, and password compromises. Nevertheless, Sarker (2016) argues for a balance between theoretical abstractions that provide *contextual specificity* and *generality*. He concurs that we should contextualize the findings by identifying relevant boundary conditions, but he also respects the generalizability of good theory, albeit without subscribing to universalism. Accordingly, we recognize the potential limitation introduced by testing our theories within the boundary conditions of our contextual abstraction.

Another limitation of using short communication messages is finding the right level of personalization. Some messages in our study may (or may not) have appeared to be more personalized, whereas others may have seemed to be more formal. Messages can also be framed to sound more threatening, friendlier, or have any other number of nuances associated with them. Previous research examined the effects of positive or negative framing of security messages on

compliance behavior and found no effect (Barlow et al., 2013). However, one should take care in interpreting differences between the strength of information, normative, and antineutralization statements based on other aspects of these messages. Future research should more fully examine the effect of how personalizing a message (vs. giving a standard, company-wide message as suggested in this paper) would affect an employee's compliance behavior. A more personalized message may make informational or normative statements more powerful.

An additional limitation is the lack of explicit statements pertaining to sanction severity and certainty in the scenarios. Such statements would allow for better control of deterrence effects. Although the scenarios indicated that penalties *will be enacted* against those individuals who violate security policies, more specific statements could be incorporated and tested to further explore the nuanced differences between persuasive information security communication and the deterrence mechanisms in place. Therefore, future studies could incorporate high/low sanction severity and certainty statements into their scenarios (see Willison et al. (2018) for example scenarios with this property), though this would effectively quadruple the number of unique scenario versions.

Another limitation is that we studied nonmalicious security policy violations that were performed by employees who were not seeking to inflict harm on their employer or to maximize their own personal gains. Previous distinctions (Guo et al., 2011; Willison & Warkentin, 2013) indicate that the motivations for such violations may be very different than for malicious violations. Future research should examine whether antineutralization communication has any effect on malicious violations. It is hard to know if antineutralization communication would influence an employee who is truly seeking to steal from his or her employer or fellow employees or to harm them in some way (Willison & Warkentin, 2013).

Finally, we used a cross-sectional design in which participants were randomly assigned to treatments. We have no measure of participants' prestudy compliance intentions, or how these may have been changed by the treatments, which is normal in laboratory research; such differences are controlled by random assignment. Future research could measure the longitudinal effects of security communication.

5.3 Implications for Future Research

This study contributes to theory by examining whether short reinforcing SETA communications containing informational, normative, and antineutralization statements are effective in diminishing information security policy violations. This study answers the call for research on persuasive communication for security training (Siponen, 2000; Thomson & von Solms, 1998). We believe there are six implications for future research.

First and foremost, our study shows the significant and meaningful effects of both neutralization techniques and the antineutralization statements designed to mitigate their effects. Prompting participants with neutralization statements increased their likelihood of violating the security policy by a similar amount as the antineutralization statements decreased them (as seen by the similar parameters in Table 2). The interaction term was not significant, which means that the effects of antineutralization are independent of deliberately provoking neutralization, likely because employees are already engaging neutralization without prompting (although prompting increases its impact). The implication is that we need more research on neutralization and antineutralization. Our research shows that a simple, one-sentence general antineutralization statement improves compliance behavior. Can the impacts of antineutralization be strengthened by increasing the length and persuasiveness of antineutralization statements? Can antineutralization be improved by targeting specific neutralization techniques or are general statements targeting all types of neutralization better? We need more research on antineutralization SETA communication.

Second, we found mixed effects from a simple one-sentence informational statement designed to explain the consequences of violating policies. Informational influence acts by providing individuals with a better understanding of the situation (the consequences of their actions in this case), which changed the balance of costs and benefits such that employees were more likely to comply with security policies. Would more information have a greater impact? For example, would explaining the reasons behind a security policy, not just the policy itself, have even stronger effects?

Third, there was no interaction between informational and antineutralization statements, indicating that they act independently of each other. Thus, the theoretical mechanisms by which the two work must be different. Using medical terminology, we would say they have different "receptors" and thus can be used in conjunction with each other much like ibuprofen and acetaminophen. We need more research on how these two act and how to better increase their joint effectiveness.

Fourth, we found that normative influence statements had no effect. Past research and theory led us to argue that SETA communications containing normative influence statements should have an effect, because it has shown internally generated normative beliefs

influence compliance behavior (Bulgurcu et al. 2010). However, in scenarios such as those we presented, where rationalization would be easy, normative statements about what the majority of employees would do was not powerful enough to influence violation intentions. The effects may have been stronger if normative influence was delivered directly from a known colleague rather than being invoked in a short general form as we did in this study (i.e., as "a recent survey of our employees"). That is, the unique reinforcing aspect of the communication we tested is different from other types of SETA programs designed to reduce security violations. We need more research to better understand why internally generated normative beliefs appear to influence behavior (at least from post hoc reports) but externally generated short normative statements provided as reinforcement prior to a decision to violate do not.

Fifth, we found some important similarities and differences between past research (which has sought to understand what factors influence security violation intentions) and our research (which strives to evaluate if SETA communications designed using that understanding can influence behavior). Past research concludes that violation decisions are influenced by information about the consequences of violating the policy and by normative beliefs about what others would do. We found that SETA communications that included informational statements about consequences of violations reduced the intention to violate, but that SETA communications incorporating normative statements about what others would do had no effect on violation intentions. This suggests that informational and normative beliefs operate in fundamentally different ways in influencing security violation intentions. This calls for more research on the differences between informational and normative influences on violation intentions, and why some factors shown to influence intentions (i.e., information) can be used to alter them whereas others (i.e., normative) cannot.

Sixth, another interesting finding of this study is in Table 4: the nonsignificant neutralization treatment (denial of injury to others) did not provide any personal benefit; whereas the significant neutralization treatment (defense of necessity) did provide personal benefits. This provides interesting insights in that cognitive dissonance may be the strongest and people may feel more motivated to use neutralizations when they perceive some direct benefit from doing so. Therefore, future research could examine a baseline condition with two other conditions: (1) a neutralization condition that offers no personal benefits and (2) a neutralization condition that offers personal benefits.

Finally, we believe that this research opens a new direction for security research. Past security research has focused on deterrence, which is appropriate given its traditional roots in criminology (D'Arcy et al., 2009). Our research shows that incorporating cognitive concepts such as informational and normative communication into information security research, as advocated by Puhakainen and Siponen (2010), has an influence on policy violations. This provides a much broader view that offers rich insights into the ways we can influence—in positive and negative ways—information systems security compliance by directly influencing the decision-making process of employees who face security-related decisions every day.

5.4 Implications for Practice

This study contributes to practice by examining specific types of persuasion that can be incorporated into organizational communication to address security vulnerabilities. First, we recommend that organizations include specific, strong informational statements that explain the purpose of information security policies when they engage in security communication. Second, we recommend that they directly communicate, in a variety of security training and communication methods, that the usage of neutralization techniques is unacceptable (i.e., antineutralization statements). Such antineutralization statements could also be personalized for different types of individual biases, much like SETA communication can be personalized to different individual traits (Johnston et al., 2016). Finally, we conclude that, even though norms can influence security behavior, including indirect normative information in short security communication is not as powerful.

6 Conclusion

This study was the first to develop and test distinct categories of persuasion techniques for information security SETA programs, including communications to employees, in common situations where employees may be tempted to rationalize insecure behavior. We found that informational statements and antineutralization statements are particularly helpful in decreasing employee intentions to violate information security policies in these situations. Our findings provide foundation for further research within the information security domain. The results also provide guidance to information security managers who design SETA programs to enhance compliance with information security policies.

References

Ajzen, I. (2002). Perceived behavioral control, self-efficacy, locus of control, and the theory of planned behavior. *Journal of Applied Social Psychology*, *32*(4), 665-683.

Akhawe, D. & Felt, A. P. (2013). Alice in Warningland: A large-scale field study of browser security warning effectiveness. In *Proceedings of the 22nd USENIX Conference on Security* (pp. 257-272). ACM

Albrechtsen, E. & Hovden, J. (2010). Improving information security awareness and behaviour through dialogue, participation and collective reflection. An intervention study. *Computers & Security*, *29*(4), 432-445.

Anderson, B. B., Vance, A., Kirwan, C. B., Eargle, D., & Jenkins, J. L. (2016). How users perceive and respond to security messages: A NeuroIS research agenda and empirical study. *European Journal of Information Systems*, *25*(4), 364-390.

Anderson, C. L. & Agarwal, R. (2010). Practicing safe computing: A multimethod empirical examination of home computer user security behavioral intentions. *MIS Quarterly*, *34*(3), 613-643.

Andrich, D. (1978). A rating formulation for ordered response categories. *Psychometrika*, *43*(4), 561-573.

Argelaguet, F., Kulik, A., Kunert, A., Andujar, C., & Froehlich, B. (2011). See-through techniques for referential awareness in collaborative virtual reality. *International Journal of Human-Computer Studies*, *69*(9), 387-400.

Aronson, E., Wilson, T. D., & Akert, A. M. (2005). *Social psychology* (5th ed.). Upper Saddle River, NJ: Prentice Hall.

Barlow, J. B., Warkentin, M., Ormond, D., & Dennis, A. R. (2013). Don't make excuses! Framing IT security training to reduce policy violation. *Computers & Security*, *39* (Part B), 145-159.

Boss, S. R., Galletta, D. F., Lowry, P. B., Moody, G. D., & Polak, P. (2015). What do systems users have to fear? Using fear appeals to engender threats and fear that motivate protective security behaviors. *MIS Quarterly*, *39*(4), 837-864.

Bulgurcu, B., Cavusoglu, H., & Benbasat, I. (2010). Information security policy compliance: An empirical study of rationality-based beliefs and information security awareness. *MIS Quarterly*, *34*(3), 523-548.

Burnkrant, R. E. & Cousineau, A. (1975). Informational and normative social influence in buyer behavior. *Journal of Consumer Research*, *2*(3), 206-215.

Burnstein, E. & Vinokur, A. (1973). Testing two classes of theories about induced shifts in individual choice. *Journal of Experimental Social Psychology*, *9*(2), 123-137.

Chatterjee, S., Sarker, S., & Valacich, J. S. (2015). The behavioral roots of information systems security: Exploring key factors related to unethical IT use. *Journal of Management Information Systems*, *31*(4), 49-87.

Ciampa, M. (2013). A comparison of password feedback mechanisms and their impact on password entropy. *Information Management & Computer Security*, *21*(5), 344-359.

Cromwell, P. & Thurman, Q. (2003). The devil made me do it: Use of neutralizations by shoplifters. *Deviant Behavior*, *24*(6), 535-550.

Crossler, R. E., Johnston, A. C., Lowry, P. B., Hu, Q., Warkentin, M., & Baskerville, R. (2013). Future directions for behavioral information security research. *Computers & Security*, *32*(1), 90-101.

D'Arcy, J., Hovav, A., & Galletta, D. (2009). User awareness of security countermeasures and its impact on information systems misuse: A deterrence approach. *Information Systems Research*, *20*(1), 79-98.

Dennis, A. R. (1996). Information exchange and use in group decision making: You can lead a group to information, but you can't make it think. *MIS Quarterly*, *20*(4), 433457.

Dennis, A. R., Hilmer, K. M., & Taylor, N. J. (1998). Information exchange and use in GSS and verbal group decision making: Effects of minority influence. *Journal of Management Information Systems*, *14*(3), 61-88.

Dhillon, G. (1999). Managing and controlling computer misuse. *Information Management & Computer Security*, *7*(4), 171-175.

EY. (2017). Path to cyber resilience: Sense, resist, react: EY's 19th global information security survey 2016-2017. Retrieved from https://www.ey.com/gl/en/industries/power---utilities/ey-the-path-to-cyber-resilience-sense-resist-react

Festinger, L. (1957). *A theory of cognitive dissonance.* Stanford, CA: Stanford University Press.

Goel, S., Williams, K., & Dincelli, E. (2017). Got phished? Internet security and human vulnerability. *Journal of the Association for Information Systems*, *18*(1), 22-44.

Greenberg, J. (1990). Employee theft as a reaction to underpayment inequity: The hidden cost of pay

cuts. *Journal of Applied Psychology*, *75*(6), 561-568.

Guo, K. H., Yuan, Y., Archer, N. P., & Connelly, C. E. (2011). Understanding nonmalicious security violations in the workplace: A composite behavior model. *Journal of Management Information Systems*, *28*(2), 203-236.

Harrington, S. J. (1996). The effect of codes of ethics and personal denial of responsibility on computer abuse judgments and intentions. *MIS Quarterly*, *20*(3), 257-278.

Henry, S. (2009). *Social deviance*. Cambridge, UK: Polity.

Herath, T. & Rao, H. R. (2009). Encouraging information security behaviors in organizations: Role of penalites, pressures and perceived effectiveness. *Decision Support Systems*, *47*(2), 154-165.

Herzog, S. (2003). The relationship between public perceptions of crime seriousness and support for plea-bargaining practices in Israel: A factorial survey approach. *The Journal of Criminal Law & Criminology*, *94*(1), 103-131.

Ifinedo, P. (2012). Understanding information systems security policy compliance: An integration of the theory of planned behavior and the protection motivation theory. *Computers & Security*, *31*(1), 83-95.

Ifinedo, P. (2014). Information systems security policy compliance: An empirical study of the effects of socialisation, influence, and cognition. *Information & Management*, *51*(69-79).

Jasso, G. (2006). Factorial survey methods for studying beliefs and judgments. *Sociological Methods & Research*, *34*(3), 334-423.

Jenkins, J. L. & Durcikova, A. (2013). What, I shouldn't have done that?: The influence of training and just-in-time reminders on secure behavior. In *Proceedings of the International Conference on Information Systems*. AIS.

Johnston, A. C. & Warkentin, M. (2010). Fear appeals and information security behaviors: An empirical study. *MIS Quarterly*, *34*(3), 549-566.

Johnston, A. C., Warkentin, M., McBride, M. E., & Carter, L. (2016). Dispositional and situational factors: Influences in IS security policy violations. *European Journal of Information Systems*, *25*(3), 231-251.

Johnston, A. C., Warkentin, M., & Siponen, M. (2015). An enhanced fear appeal framework: Leveraging threats to the human asset through sanctioning rhetoric. *MIS Quarterly*, *39*(1), 113-134.

Kaplan, M. F. & Miller, C. E. (1987). Group decision making and normative versus informational influence: Effects of type of issue and assigned decision rule. Journal of *Personality & Social Psychology*, *53*(2), 306-313.

Karjalainen, M., & Siponen, M. (2011). Toward a new meta-theory for designing information systems (IS) security training approaches. *Journal of the Association for Information Systems*, *12*(8), 518-555.

Kerlinger, F. (1973). *Foundations of behavioral research* (2nd ed.). London: Holt Reinhart & Winston.

Kim, H.-W. & Kankanhalli, A. (2009). Investigating user resistance to information systems implementation: A status quo bias perspective. *MIS Quarterly*, 33(3), 567-582.

Lanza, M. L. (1988). Technical notes—Development of a vignette: A data collection instrument about patient assault. *Western Journal of Nursing Research*, *10*(3), 346-351.

Latané, B. (1981). The psychology of social impact. *American Psychologist*, *36*(4), 343-356.

Lauder, W. (2002). Factorial survey methods: A valuable but under-utilized research method in nursing research? *Nursing Times Research*, *7*(1), 35-43.

Lebek, B., Uffen, J., Breitner, M. H., Neumann, M., & Hohler, B. (2013). Employees' information security awareness and behavior: A literature review. In *Proceedings of the 46th Hawaii International Conference on System Sciences* (pp. 2978-2987). AIS.

Lewis, W., Agarwal, R., & Sambamurthy, V. (2003). Sources of influence on beliefs about information technology use: An empirical study of knowledge workers. *MIS Quarterly*, *27*(4), 657-678.

Liang, H., Saraf, N., Hu, Q., & Xue, Y. (2007). Assimilation of enterprise systems: The effect of institutional pressures and the mediating role of top management. *MIS Quarterly*, *31*(1), 59-87.

Liang, H. & Xue, Y. (2009). Avoidance of information technology threats: A theoretical perspective. *MIS Quarterly*, *33*(1), 71-90.

Liang, H. & Xue, Y. (2010). Understanding security behaviors in personal computer usage: A threat avoidance perspective. *Journal of the Association for Information Systems*, *11*(7), 394-413.

Lim, V. K. G. (2002). The IT way of loafing on the job: Cyberloafing, neutralizing and organizational justice. *Journal of Organizational Behavior*, *23*(5), 675-694.

Lowry, P. B. & Moody, G. D. (2014). Proposing the control-reactance compliance model (CRCM) to

explain opposing motivations to comply with organisational information security policies. *Information Systems Journal, 25*(5), 433-463.

Minor, W. W. (1981). Techniques of neutralization: A reconceptualization and empirical examination. *Journal of Research in Crime and Delinquency, 18*(2), 295-318.

Myyry, L., Siponen, M., Pahnila, S., Vartiainen, T., & Vance, A. (2009). What levels of moral reasoning and values explain adherence to information security rules? An empirical study. *European Journal of Information Systems, 18*(2), 126-139.

O'Fallon, M. & Butterfield, K. (2005). A review of the empirical ethical decision-making literature: 1996-2003. *Journal of Business Ethics*, *59*(4), 375-413.

Oppenheimer, D. M., Meyvis, T., & Davidenko, N. (2009). Instructional manipulation checks: Detecting satisficing to increase statistical power. *Journal of Experimental Social Psychology, 45*(4), 867-872.

Paternoster, R. & Simpson, S. (1996). Sanction threats and appeals to morality: Testing a rational choice model of corporate crime. *Law & Society Review, 30*(3), 549-584.

Pershing, J. L. (2003). To snitch or not to snitch? Applying the concept of neutralization techniques to the enforcement of occupational misconduct. *Sociological Perspectives*, *46*(2), 149-178.

Petty, R. E. & Cacioppo, J. T. (1986). *Communication and persuasio*n. New York, NY: Springer.

Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, *88*(5), 879-903.

Posey, C., Roberts, T. L., & Lowry, P. B. (2015). The impact of organizational commitment on insiders' motivation *Journal of Management Information Systems* to protect organizational information assets, *32*(4), 179-214.

Post, G. V. & Kagan, A. (2007). Evaluating information security tradeoffs: Restricting access can interfere with user tasks. *Computers & Security, 26*(3), 229-237.

Puhakainen, P. & Siponen, M. (2010). Improving employees' compliance through information systems security training: An action research study. *MIS Quarterly, 34*(4), 757-788.

Ransbotham, S. & Mitra, S. (2009). Choice and chance: A conceptual model of paths to information security compromise. *Information Systems Research*, *20*(1), 121-139.

Rennie, L. (1982). Research note: Detecting a response set to Likert-style attitude items with the rating model. *Educational Research and Perspectives*, *9*(1), 114-118.

Rossi, P. H. & Anderson, A. B. (1982). The factorial survey approach: An introduction, In P. H. Rossi & S. L. Nock (Eds.), *Measuring social judgments: The factorial survey approach* (pp. 15-67). Beverly Hills, CA: Sage.

Sarker, S. (2016). Building on Davison and Martinsons' concerns: A call for balance between contextual specificity and generality in IS research. *Journal of Information Technology*, *31*(3), 250-253.

Seron, C., Pereira, J., & Kovath, J. (2006). How citizens assess just punishment for police misconduct. *Criminology, 44*(4), 925-960.

Shaw, M. (1981). *Group dynamics*: *The psychology of small group behavior* (3rd ed.). New York, NY: McGraw Hill.

Shlay, A. B., Tran, H., Weinraub, M., & Harmon, M. (2005). Teasing apart the child care conundrum: A factorial survey analysis of perceptions of child care quality, fair market price and willingness to pay by low-income, African American parents. *Early Childhood Research Quarterly*, *20*(4), 393-416.

Silic, M., Barlow, J. B., & Back, A. (2017). A new perspective on neutralization and deterrence: Predicting shadow IT usage. *Information & Management, 54*(8), 1023-1037.

Siponen, M. (2000). A conceptual foundation for organizational information security awareness. *Information Management & Computer Security*, *8*(1), 31-41.

Siponen, M., Mahmood, M. A., & Pahnila, S. (2014). Employees' adherence to information security policies: An exploratory field study. *Information & Management, 51*(2), 217-224.

Siponen, M. & Vance, A. (2010). Neutralization: New insights into the problem of employee information systems security policy violations. *MIS Quarterly*, *34*(3), 487-502.

Siponen, M., Vance, A., & Willison, R. (2012). New insights into the problem of software piracy: The effects of neutralization, shame, and moral beliefs. *Information & Management*, *49*(7/8), 334-341.

Son, J.-Y. (2011). Out of fear or desire? Toward a better understanding of employees' motivation to follow IS security practices. *Information & Management*, *48*(7), 296302.

Straub, D., Boudreau, M.-C., & Gefen, D. (2004). Validation guidelines for IS positivist research. *Communications of the AIS*, *13*(1), 381-427.

Straub, D. W. & Welke, R. J. (1998). Coping with systems risk: Security planning models for management decision making. *MIS Quarterly*, *22*(4), 441-469.

Subramani, M. R. & Rajagopalan, B. (2003). Knowledge-sharing and influence in online social networks via viral marketing. *Communications of the ACM*, *46*(12), 300-307.

Sussman, S. W. & Siegal, W. S. (2003). Informational influence in organizations: An integrated approach to knowledge adoption. *Information Systems Research*, *14*(1), 47-65.

Sykes, G. & Matza, D. (1957). Techniques of neutralization: A theory of delinquency. *American Sociological Review*, *22*(6), 664-670.

Thaler, R. & Sunstein, C. (2008). *Nudge: Improving decisions about health, wealth, and happiness.* New Haven, CT: Yale University Press.

Thomson, M. E. & von Solms, R. (1998). Information security awareness: Educating your users effectively. *Information Management & Computer Security*, *6*(4), 167-173.

Trevino, L. K. (1992). Experimental approaches to studying ethical-unethical behavior in organizations. *Business Ethics Quarterly*, *2*(2), 121-136.

Trinkle, B. S., Crossler, R. E., & Warkentin, M. (2014). I'm game, are you? Reducing realworld security threats by managing employee activity in virtual environments. *Journal of Information Systems*, *28*(2), 307-327.

Tsohou, A., Karyda, M., & Kokolakis, S. (2015). Analyzing the role of cognitive and cultural biases in the internalization of information security policies: Recommendations for information security awareness programs. *Computers & Security*, *52*, 128-141.

Tversky, A. & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, *211*(4481), 453-458.

Ur, B., Kelley, P. G., Komanduri, S., Lee, J., Maass, M., Mazurek, M. L., et al. (2012). How does your password measure up? The effect of strength meters on password creation. In *Proceedings of the 21st USENIX Security Symposium* (pp. 65-80). ACM.

Vance, A., Anderson, B., Kirwan, B., Eargle, D. (2014). Using measures of risk perception to predict information security behavior: Insights from electroencephalography (EEG), *Journal of the Association for Information Systems*, *15*(10), 679-722.

Vance, A., Lowry, P. B., & Eggett, D. (2015). Increasing accountability through user interface design artifacts: A new approach to address the problem of access-policy violations. *MIS Quarterly*, *39*(2), 345-366.

Wall, J. D., Lowry, P. B., & Barlow, J. B. (2016). Organizational violations of externally governed privacy and security rules: Explaining and predicting selective violations under conditions of strain and excess. *Journal of the Association for Information Systems*, *17*(1), 39-76.

Wallander, L. (2009). 25 years of factorial surveys in sociology: A review. *Social Science Research*, *38*(3), 505-520.

Warkentin, M., Willison, R., & Johnston, A. C. (2011). The role of perceptions of organizational injustice and techniques of neutralization in forming computer abuse intentions. In *Proceedings of the 17th Americas Conference on Information Systems*. AIS.

Weber, J. (1992). Scenarios in business ethics research: Review, critical assessment, and recommendations. *Business Ethics Quarterly*, *2*(2), 137-160.

Willison, R. & Warkentin, M. (2010). The expanded security action cycle: A temporal analysis "left of bang." In *Proceedings of the Dewald Roode Information Security Workshop, IFIP WG8.11/11.13.* IOS.

Willison, R. & Warkentin, M. (2013). Beyond deterrence: An expanded view of employee computer abuse. *MIS Quarterly*, *37*(1), 1-20.

Willison, R., Warkentin, M., & Johnston, A. C. (2018) Examining employee computer abuse intentions: Insights from justice, deterrence and neutralization perspectives. *Information Systems Journal*, *28*(2), 266-293.

Xue, Y., Liang, H., & Wu, L. (2011). Punishment, justice, and compliance in mandatory IT settings. *Information Systems Research*, *22*(2), 400-414.

Zitek, E. M. & Hebl, M. R. (2007). The role of social norm clarity in the influenced expression of prejudice over time. *Journal of Experimental Social Psychology*, *43*(6), 867-876.

Appendix A. Scenarios

Table A1. Baseline Scenario

At Crossroads, Inc.*, management has been focusing on increasing compliance with IT security policies. The company has penalties in place for employees who violate policies. The company recently developed an IT security training program where employees read information about IT security policies and have a group discussion about them. Here is an excerpt from the security training program materials:

"As stated in our security policies, employees should not share computer passwords with other coworkers. This applies equally to all employees.

"[Insert normative influence statement here.]"

"[Insert antineutralization influence statement here.]"

"[*Insert informational influence statement here*.]"

Matt* is one of the employees at Crossroads, Inc. who has completed the security training program. While out of town, Matt gets an urgent call from a coworker, John. John tells Matt that he has to get a project done right away to meet a deadline, but he needs some information from Matt. Matt recalls that the information John needs is saved on the hard drive of Matt's office computer, which is not set up for remote access. John asks Matt to share his password in order to access the needed information for his report.

[**Insert neutralization technique here**.] Matt decides to go ahead and share his password with John.

Normative influence statement: A recent survey of our employees concerning this policy showed that over 85 percent would not share their password, even with another employee, regardless of the circumstances.

(No normative influence statement)

Antineutralization statement: Even though people believe that sharing passwords can be justified under certain circumstances without any real consequences, adherence to this policy is important; sharing of passwords should not be justified for any reason.

(No antineutralization influence statement)

Informational influence statement: While it may not appear to be the case, there are often real consequences of sharing passwords that extend beyond the person disobeying the policy, such as when a company recently experienced unauthorized access to confidential customer information because an employee shared his password.

(No informational influence statement)

Denial of injury: Matt knows that John is trustworthy and feels that no harm would result from sharing his password with John this one time. Besides, he can change his password afterwards.

Defense of necessity: Matt knows that John's project is critical to the success of their department. If the project fails, there will be consequences not only for John, but also for Matt. Matt is unable to get to the office today, so he feels there is no other choice.

(No neutralization technique)

Note: There are 2x2x2x3 (24) versions of this scenario.

* Each scenario has unique company and individual names. We reviewed and changed names during multiple rounds of expert review. The list of company and individual names is available upon request.

Table A2. Example Scenario

At Crossroads, Inc., management has been focusing on increasing compliance with IT security policies. The company has penalties in place for employees who violate policies. The company recently developed an IT security training program where employees read information about IT security policies and have a group discussion about them. Here is an excerpt from the security training program materials:

"As stated in our security policies, employees should not share computer passwords with other coworkers. This applies equally to all employees.

"A recent survey of our employees concerning this policy showed that over 85 percent would not share their password, even with another employee, regardless of the circumstances."

"Even though people believe that sharing passwords can be justified under certain circumstances without any real consequences, adherence to this policy is important; sharing of passwords should not be justified for any reason."

"While it may not appear to be the case, there are often real consequences of sharing passwords that extend beyond the person disobeying the policy, such as when a company recently experienced unauthorized access to confidential customer information because an employee shared his password."

Matt* is one of the employees at Crossroads, Inc. who has completed the security training program. While out of town, Matt gets an urgent call from a coworker, John. John tells Matt that he has to get a project done right away to meet a deadline, but he needs some information from Matt. Matt recalls that the information John needs is saved on the hard drive of Matt's office computer, which is not set up for remote access. John asks Matt to share his password in order to access the needed information for his report.

Matt knows that John is trustworthy and feels that no harm would result from sharing his password with John this one time. Besides, he can change his password afterwards. Matt decides to go ahead and share his password with John

Note: Formatting in the example scenarios in Table A1 and A2 indicate the following: antineutralization, **normative influence**, *informational influence*, ***denial of injury.***

Appendix B. Survey Measures

Filter Questions
Have you held a job in a workplace that had guidelines, work rules, or policies for employees? YES/NO
Have you held a job in which you used a computer for your work? YES/NO
[If participants answer "no" to either, the survey ends and Qualtrics does not collect further data]

Manipulation Check
Please select an answer for the following items as they relate to the scenario above:
[Each question only present when the corresponding statement was included in the scenario.]

In **this** scenario, the security awareness training material clearly states that:
a. employees should never justify sharing passwords.
b. employees will receive written warnings for sharing passwords.

According to **this** scenario, the security awareness training material includes:
a. a statement from the CEO about the importance of adhering to the policy.
b. a summary of results of a recent employee survey concerning the policy.

According to **this** scenario, the security awareness training material includes:
a. a description of possible consequences—other than penalties to the employee—of sharing passwords.
b. a summary of other IT security policies related to the password policy.

How does Matt justify sharing his password in **this** scenario?
a. He believes that no harm will result from sharing his password.
b. He believes that sharing his password is critical to the success of his department.
c. He believes that because he has been a good employee for many years he can share his password.

[General question for the one "baseline" scenario that had no manipulation statements]
In **this** scenario, the security awareness training material clearly states that:
a. employees will receive written warnings for sharing passwords.
b. the IT security policy in question applies equally to all employees.

Content validity (realism check)	SD D N A SA
I could imagine a similar scenario taking place at my company.	1 2 3 4 5

Dependent variable measures (behavioral intention)	SD D N A SA
In this situation, I would do the same as Matt.	1 2 3 4 5
If I were Matt, I would have also shared my password.	1 2 3 4 5
I think I would do what Matt did.	1 2 3 4 5
I think others would do the same if they were Matt.	1 2 3 4 5

Demographic items

I am	Male / Female
My age is	(freeform integer)
Years of work experience:	(freeform integer)
Highest education:	Some high school / High school / Undergraduate / Graduate

Appendix C. Data Filtering and Robustness Checks

Data Filtering

Considering the difficulty in obtaining reliable data from online survey participants, we wanted to ensure that we received quality responses for our data analysis. Therefore, we contracted with Qualtrics to acquire survey responses from 200 participants who had passed all filter, response set, and manipulation check questions in our design. Each time a recruited participant answered one of these questions unsatisfactorily, the survey ended early for that participant, and the response was marked incomplete. This process continued until 200 completed surveys met these criteria.

First, we filtered participants to include only those who worked full-time in companies with computer policies. Of the 908 initial participants, 262 did not pass these filter questions and saw no scenarios, resulting in 646 participants who completed at least one scenario. When participants failed the response set question, indicating they were not paying attention to the survey questions, the survey ended and these participants were not included in the final data set (Andrich, 1978; Kerlinger, 1973; Rennie, 1982). This included 102 participants, reducing our data set to 544 participants who passed all filter or response set questions. Based on the correlation of failing manipulation checks with lack of motivation and disregard for research instructions, experimental researchers recommend that data including failed manipulation checks not be included in data analysis (Oppenheimer et al., 2009). Of the remaining 544 participants, 303 did not correctly answer all manipulation check questions, leaving 241 participants with valid responses to at least one scenario. However, of these 241, 41 closed the survey early, not completing the demographic portion of the survey and in some cases only completing one of the two scenarios.

Additional Data Analysis as Robustness Checks

Our main data analysis considered the 200 fully complete and valid survey completions (400 total scenario responses) that we contracted for and received from Qualtrics. However, Qualtrics kindly provided the partial data from all participants who started the survey. Given that a large portion of the data were removed due to incomplete responses or failed manipulation checks, we repeated our analysis summarized in Table 2 of the main text on three additional data sets that included excluded participants.

The first data set we used in our robustness checks ("A") consisted of all valid survey responses where filter, response set, and manipulation check questions were correct (241 participants; 451 scenarios), even though some of these had missing demographic data because the survey was closed early. In the second data set ("B"), we only excluded the participants who did not pass the filter questions and the response set questions, keeping incomplete surveys where in some cases the manipulation check was answered incorrectly (544 participants; 831 scenarios). The third data set used for robustness analysis ("C") consisted of all data where participants passed the initial filter questions, regardless of whether they answered response set and/or manipulation check questions correctly (646 subjects, 954 scenarios). Given that only the 200 final participants completed demographic information, the demographic variables could not be examined as control variables in these three data sets.

The results of analyzing these data sets were largely similar to the results of the main analysis. In particular, in all data sets, neutralization statements increased the likelihood that a participant would state intentions to violate the policy, antineutralization statements were consistently powerful in reducing stated intentions to violate, and there was an effect order. There were two differences in the results between these additional data sets and the main analysis. First, realism was significant in these data sets—those who viewed the scenarios as more realistic were more likely to state intentions to violate the policy. Second, information influence statements were only statistically significant in data set "B" (further reflecting a slight effect that is statistically significant in some data subsets, but not others). See Table C1 for summarized results.

In these additional data sets, we also controlled for the effects of completing the survey, passing manipulation checks, and passing the response set questions. Of these, only the responses to the response set questions were significantly correlated with stating intentions to violate policy. Specifically, those who passed the response set questions, indicating they paid attention to the survey, were less likely to state intentions of violating a policy. This is intuitive, because those not paying attention would be more likely to select "agree" or "strongly agree" on most questions, including questions of whether they would violate a policy.

Table C1. Robustness Check Results

	Data set "A"		Data set "B"		Data set "C"	
	Estimate	p	Estimate	p	Estimate	p
Intercept	1.816	<0.001	1.468	<0.001	1.338	<0.001
Use of neutralization techniques	0.304**	0.005	0.333***	<0.001	0.305***	<0.001
Informational influence statement	-0.151	0.187	-0.182*	0.019	-0.102	0.147
Normative influence statement	0.101	0.324	0.055	0.447	0.000	0.998
Antineutralization influence statement	0.377***	<0.001	-0.294***	<0.001	-0.294***	<0.001
Order	-0.121*	0.039	-0.167**	0.005	-0.199***	<0.001
Realism	0.088†	0.094	0.179***	<0.001	0.224***	<0.001
Participant fully completed the survey	-0.124	0.501	-0.017	0.897	-0.064	0.596
Participant passed manipulation checks			0.155	0.146	-0.070	0.455
Participant passed response set					-0.414***	<0.001
Note : †$p < 0.10$; *$p < 0.05$; **p <0.01; *** $p < 0.01$						

Finally, we also performed a robustness check where we analyzed only the responses to the first scenario seen by each participant. These results are shown in Table C2.

Table C2. Robustness Check for First Scenarios

	Full results (n = 400)		Only first scenario n = 200)	
	Estimate	p	Estimate	P
Intercept	-0.443	0.216	-0.349	0.425
Use of neutralization techniques	0.227*	0.026	0.193	0.228
Informational influence statement	-0.175†	0.092	-0.188	0.201
Normative influence statement	0.065	0.216	0.018	0.902
Antineutralization influence statement	-0.320**	0.001	-0.429**	0.003
Order	-0.142*	0.010	n/a	n/a
Realism	0.058	0.267	0.040	0.526
Gender (Female)	0.092	0.465	0.154	0.296
Age	0.000	0.982	0.001	0.962
Work experience	-0.005	0.624	-0.006	0.587
Education	0.097	0.558	0.073	0.698
Note : †$p < 0.10$; *$p < 0.05$; **p <0.01; *** $p < 0.01$;	n = 400 scenarios		n = 200 scenario	

As shown in the table, there were some small differences when testing only the first scenarios. Specifically, the explicit use of neutralization techniques in the scenario was no longer statistically significantly related to intentions to violate the policy. Informational statements, which were significant at the $p < 0.1$ level in the main results, did not have a statistically significant effect when analyzing only the first scenarios. However, it should be noted that the coefficient is actually greater than in the main results, but cutting the sample size in half greatly increased the *p* value. That is, the effect size is similar, but the statistical significance is weak.

Appendix D. Raw Data for Intention Scores by Group

Table D1. Intention Scores by Group

	Present	Not Present
Antineutralization statement	2.420	2.790
Information influence statement	2.487	2.692
Normative influence statement	2.608	2.593
Denial of injury neutralization statement	2.486	2.448
Defense of necessity neutralization statement	2.849	

Appendix E. Interaction Effects

We analyzed all interaction effects between neutralization types and the hypothesized effects of informational, normative, and antineutralization statements and found no significant results (see Table D1). In other words, the effects of the different types of persuasion were similar regardless of the specific neutralization type used in the various scenarios.

Table E1. Interaction Effect Tests

	Both neutralization types		Denial of injury		Defense of necessity	
	Estimate	p	Estimate	p	Estimate	p
Intercept	-0.464	0.210	-0.565	0.147	-0.834	0.080
Use of neutralization techniques	0.292	0.158	0.039	0.878	0.480*	0.028
Informational influence statement	-0.189	0.275	-0.174	0.324	-0.164	0.340
Normative influence statement	0.091	0.607	0.086	0.629	0.080	0.652
Antineutralization influence statement	-0.258†	0.088	-0.252†	0.093	-0.255	0.105
Order	-0.140*	0.11	0.080	0.344	-0.071	0.765
Realism	0.058	0.267	0.110†	0.060	0.091	0.136
Gender (female)	0.092	0.471	0.178	0.212	0.008	0.957
Age	0.000	0.974	-0.002	0.862	0.009	0.460
Work experience	-0.005	0.617	-0.004	0.723	-0.009	0.413
Education	0.094	0.567	0.162	0.361	0.055	0.765
Note : †$p < 0.10$; *$p < 0.05$; **$p < 0.01$; *** $p < 0.01$						

Appendix F. SPSS Syntax

```
GENLINMIXED
 /DATA_STRUCTURE SUBJECTS=Subject REPEATED_MEASURES=Order
COVARIANCE_TYPE=DIAGONAL
 /FIELDS TARGET=DV_factor TRIALS=NONE OFFSET=NONE
 /TARGET_OPTIONS DISTRIBUTION=NORMAL LINK=IDENTITY
 /FIXED EFFECTS=Neutrcombined1anytechniquepresent Infostatementpresent
Normstatementpresent AntiNeutrstatementpresent Order Realism GenderM1
Age WorkExp Education USE_INTERCEPT=TRUE
 /BUILD_OPTIONS TARGET_CATEGORY_ORDER=ASCENDING
INPUTS_CATEGORY_ORDER=ASCENDING MAX_ITERATIONS=100 CONFIDENCE_LEVEL=95
DF_METHOD=RESIDUAL COVB=ROBUST PCONVERGE=0.000001(ABSOLUTE) SCORING=0
SINGULAR=0.000000000001
 /EMMEANS_OPTIONS SCALE=ORIGINAL PADJUST=LSD.
```

About the Authors

Jordan B. Barlow is an assistant professor in the Department of Graduate Programs in Software at the University of St. Thomas. His research, which focuses on behavioral aspects of computer-mediated collaboration and information systems security, has been published in *MIS Quarterly*, *Journal of Management Information Systems*, *Journal of the Association for Information Systems*, *Information & Management*, and other high-quality journals and conferences.

Merrill Warkentin is the James J. Rouse Endowed Professor of Information Systems in the College of Business at Mississippi State University. His research, which focuses primarily on IS security and privacy behaviors, has appeared in *MIS Quarterly*, *Journal of MIS*, *Journal of the Association for Information Systems*, *European Journal of Information Systems*, *Information Systems Journal*, *Decision Sciences*, and others. He holds or has held editorial positions at *MIS Quarterly*, *Information Systems Research*, *Journal of the Association for Information Systems*, *European Journal of Information Systems*, *Decision Sciences*, *Information & Management*, and others.

Dustin Ormond is an assistant professor of business intelligence and analytics at Creighton University. His research, which primarily focuses on behavioral information security, affective computing, and deception, has appeared in *Journal of the Association for Information Systems*, *Computers & Security*, *Journal of Information Privacy and Security*, *Journal of Computer Information Systems*, and others.

Alan R. Dennis is a professor of information systems and holds the John T. Chambers Chair of Internet Systems in the Kelley School of Business at Indiana University. He was named a fellow of the Association for Information Systems in 2012. His research focuses on three main themes: team collaboration, fake news on social media, and information security. He is president elect of the Association for Information Systems.