

Training to Mitigate Phishing Attacks Using Mindfulness Techniques

MATTHEW L. JENSEN, MICHAEL DINGER, RYAN T. WRIGHT, AND JASON BENNETT THATCHER

MATTHEW L. JENSEN (mjensen@ou.edu; corresponding author) is an associate professor of management information systems and a co-director of the Center for Applied Social Research at the University of Oklahoma. His interests include computer-aided decision making, human-computer interaction, and computer-mediated communication. He studies how people attribute credibility in mediated interactions and how people filter and evaluate information they find online. His research has been published in *Information Systems Research*, *Journal of Management Information Systems*, *MIS Quarterly*, and other journals. He has been primary investigator or co-primary investigator on externally funded research projects totaling more than \$8 million.

MICHAEL DINGER (mdinger@uscupstate.edu) is an assistant professor of management in the Johnson College of Business and Economics at the University of South Carolina Upstate. He received a Ph.D. in management information systems from Clemson University. His research interests include IT workforce management, information security, and absorptive capacity. His work appears in *MIS Quarterly*, *Information Systems Research*, and *Journal of the Association for Information Systems*.

RYAN T. WRIGHT (rtwright@virginia.edu) is an associate professor in the McIntire School of Commerce at the University of Virginia. He holds a Ph.D. from Washington State University and a BS and MBA from the University of Montana. His research interests include IT security and privacy, diffusion of innovations, and digital commerce. His research on cybersecurity has been funded by the State of Massachusetts and the National Science Foundation.

JASON BENNETT THATCHER (jthatch@clemson.edu) is a professor of information systems at Clemson University. He also holds a faculty appointment at the Information Technology University-Copenhagen. His research examines the influence of individual beliefs and characteristics on information technology use. He also studies strategic and human resource management issues related to the application of information technologies in organizations. His work appears in *MIS Quarterly*, *Journal of Applied Psychology*, and other journals. He serves as a senior editor at *MIS Quarterly* as well as president of the Association for Information Systems.

ABSTRACT: Phishing attacks are at a record high and are causing billions of dollars in losses. To mitigate phishing's impact, organizations often use rule-based training to teach individuals to identify certain cues or apply a set of rules to avoid phishing attacks. The rule-based approach has improved organizational defenses against phishing; however, regular repetition of rule-based training may not yield increasing resistance to attacks. To expand the toolkit available to combat phishing attacks, we

used mindfulness theory to develop a novel training approach that can be performed after individuals are familiar with rule-based training. The mindfulness approach teaches individuals to dynamically allocate attention during message evaluation, increase awareness of context, and forestall judgment of suspicious messages—techniques that are critical to detecting phishing attacks in organizational settings, but are unaddressed in rule-based instruction. To evaluate the efficacy of our approach, we compared rule-based and mindfulness training programs in a field study at a U.S. university that involved 355 students, faculty, and staff who were familiar with phishing attacks and received regular rule-based guidance. To evaluate the robustness of the training, we delivered each program in text-only or text-plus-graphics formats. Ten days later, we conducted a phishing attack on participants that used both generic and customized phishing messages. We found that participants who received mindfulness training were better able to avoid the phishing attack. In particular, improvement was observed for participants who were already confident in their detection ability and those who reported low e-mail mindfulness and low perceptions of Internet risk. This work introduces and provides evidence supporting a new approach that may be used to develop anti-phishing training.

KEY WORDS AND PHRASES: information security, mindfulness, mindlessness, phishing, security training, signal detection.

We all receive them: e-mails that ask us to click on a link to avoid catastrophes such as immediate closure of our bank accounts or imminent deletion of our e-mail. Some messages are readily identified as fraudulent, but others might require more consideration, and a few fool even the keenest evaluators [33]. These messages come from cybercriminals known as *phishers* who use social engineering techniques to imitate electronic communications from a trustworthy source and steal credentials, collect private information, or install malicious software [31, 57]. For example, a common phishing attack consists of sending malicious, unsolicited e-mails that mirror legitimate messages. If a targeted individual falls for a phishing attack by exposing sensitive information or allowing the installation of malware, the phisher may then sell the stolen information to other criminals or use it to assume a false identity, commit financial fraud, or steal additional private information (e.g., from the individual's organization). Phishing messages may mimic legitimate messages about activities and practices within an organization, online transactions, or social exchanges, and they request recipients to perform some action (e.g., log in to a fabricated website, provide private information, or download malware). Phishers operate with communication artifacts, technologies, and processes that are ingrained in organizations and fundamental to interconnected individuals and organizations (see [60]), and phishers gain compliance using persuasion techniques such as liking, time scarcity, and social proof [91].

According to the Anti-Phishing Working Group, an international security industry coalition, the number of phishing attacks leveled at individuals and organizations increased substantially in the first quarter of 2016 compared with prior years [5]. Data belonging to millions of individuals have been stolen and popular organizations

such as banks, retailers, and service providers have been victimized in phishing attacks. In fact, the U.S. Federal Bureau of Investigation posted a warning in April 2016 that it received reports from more than 17,000 victims, which accounted for \$2.3 billion in losses [19]. Beyond financial losses, damages from phishing attacks can also include reputational harm (e.g., Sony [28]), theft of corporate secrets (e.g., LinkedIn [41]), and exposure of classified information (e.g., the White House [62]). The rates at which individuals fall victim to phishing attacks are also troubling. A simulated phishing attack on over 10,000 people in a university successfully captured credentials from roughly 9 percent of students and alumni and 5 percent of faculty and staff [56]. Furthermore, security experts predict that the threat from similar phishing attacks will continue to grow in reach and sophistication [38]. Even more troubling, industry estimates suggest that 91 percent of *all* cyber attacks begin with a phishing attack [70].

To prevent phishing attacks, organizations often rely on three techniques: (1) automated removal or quarantine of phishing messages and corresponding websites; (2) automated warning mechanisms that notify individuals when they encounter a suspicious message or website; and (3) behavioral training during which individuals are taught to identify and report attacks [31]. Automated techniques (quarantine and warning mechanisms) have been widely investigated (see [1, 44]) and have improved defenses against phishing attacks [2]. Yet, despite the protections they offer, automated techniques have a limited ability to detect new permutations of phishing attacks [29]. Phishing attacks have changed substantially from spam-like messages to focused attacks on individuals [77]. In 2014, security organizations estimated that a phishing attack targeted an average of only 23 individuals [78]. To mitigate losses from such targeted attacks, researchers have suggested a focus on end-user precautions (e.g., behavioral training) [64]. Behavioral training not only helps organization members recognize and avoid phishing messages but also engages them in identifying novel attacks as input for refining automated anti-phishing techniques.

Security researchers and practitioners have developed numerous behavioral training programs to teach individuals to identify phishing attacks. Most of these training programs share a similar *rule-based* approach, wherein trainers build rules to guide the user's reaction. Rule-based training approaches frame phishing mitigation as an identification task. Grounded in signal detection theory [24], this approach trains individuals to identify certain cues (e.g., [44]) and take protective action [68]. To the extent individuals identify the cues and follow the rules (e.g., do not click on a link in an e-mail from an unknown sender; look for HTTPS in the address bar), they will be protected from phishing attacks. The rule-based approach to training has been extensively studied, and empirical evidence has confirmed its ability to increase resistance to phishing attacks [14, 44]. Through repetition of training and reminders, the rule-based approach assists individuals to internalize and habitually apply behaviors [6]. For example, organizations often send out reminders or require training regarding safe e-mail practices on an annual or semiannual basis. Across government, business, and community organizations, such training forms an important layer of organizational defense against phishing attacks [26].¹

However, repetition of rule-based training may not yield increasing resistance to phishing attacks for several reasons. First, repeated rule-based training may lose marginal effectiveness because it can result in a (genuine or false) sense of mastery over training concepts. Effective security training takes into account individuals' current level of knowledge [67]. As training is repeated, individuals develop a sense of familiarity with training concepts and increasingly *think* they already understand them. Therefore, to the individual, additional attention to repeated training may appear to be unnecessary. Second, though rule-based training may result in habitual use of predetermined cues and rules, such rules may not provide adequate protection against customized or novel phishing attacks that deviate from those rules. Past research on habit and system use suggests that changes in behavior do not necessarily correspond to changes in systems or environment [66]. Third, rule-based training ignores the limits of human cognition and neglects the allocation of attention, a critical aspect of message evaluation that makes phishing attacks especially difficult to detect. When a new message arrives, determining whether it is a threat is often an ancillary task that is superseded by more operationally relevant tasks (e.g., responding quickly). Furthermore, if the expectations of secure behavior are overly complex, individual compliance will be less likely [15]. Consequently, phishers successfully exploit distraction or routinized, mindless responses that stem from such divided attention [17, 86].

To enhance behavioral training, approaches are needed that train individuals to allocate attention to customized attacks in organizational environments. While signal detection theory and rule-based approaches to training are well-suited to creating a basic understanding of phishing, they may not always provide the dynamic awareness necessary for resistance to evolving and customized phishing attacks. We argue that once individuals have a basic understanding of how to detect and protect themselves from phishing attacks, the training approach and underlying theory guiding the approach should change to increase resistance. Therefore, we pose the following research question: *Does supplemental training using a mindfulness approach improve resistance to phishing attacks more than additional training using the rule-based approach?* Such research is necessary because, although individuals constitute the final defense against phishing attacks, scant academic attention has been paid to the comparative design of anti-phishing training programs [44].

To answer the research question, we created two computer-based training programs using different underlying theories. The first relied on signal detection theory to create rule-based training similar to what is used widely in practice (e.g., [14, 44]). The second relied on attention and mindfulness theory [46, 47] to instruct individuals on mindfulness techniques that can be applied during message evaluation. We implemented these training approaches in text-only and text-plus-graphics formats and tested them in a field study at a Midwestern university where phishing attacks were a regular occurrence and individuals received rule-based training once or twice a semester via warnings and guidance on how to recognize phishing messages. To judge the effectiveness of the training approaches, we conducted a phishing attack at the university. With the cooperation of university officials, we distributed generic

and customized phishing messages to students, faculty, and staff that solicited their university credentials. We evaluated the effectiveness of the anti-phishing training approaches by examining which training reduced the proportion of participants who responded to the phishing attack.

Theoretical Development

Nearly all organizational members receive a steady flow of messages that traverse digital channels (e.g., e-mail, chat, text messages, or social media). Individuals must balance multiple goals when processing these messages. They must manage and properly respond to the near constant influx of messages. They must also complete work-related tasks quickly and efficiently. In attempting to accomplish these tasks, individuals expend finite cognitive resources. Managing multiple pressures (e.g., rapidly processing information and completing operational tasks) can lead to failure by workers to prioritize organizational and individual security practices. Research has found that individuals frequently rely on simple decision rules, mental shortcuts, or habit when taking protective actions (e.g., [83]) and rarely consider or even suppress consideration of security risks when warnings are raised [4].

For organization leaders seeking to mitigate phishing attacks, organization members' reliance on simple decision rules and mental shortcuts can be problematic [17, 86]. Phishers appear to understand this tendency and compose phishing messages that exploit (e.g., by using message designs that appear legitimate), or messages that encourage cursory message processing (e.g., by inducing time pressure). Such tactics have been shown to be highly effective in soliciting responses to phishing attacks [17, 86]. Conversely, research investigating individuals who successfully identified phishing attacks revealed they did not solely rely on simple decision rules, but rather allocated attention to suspicious messages and were alert to how the suspicious messages aligned with their expectations, work tasks, policies of proper e-mail usage, and other contextual factors [90].

We drew on these findings to develop a novel training approach, which teaches individuals to construct new strategies and mental models for allocating attention while processing messages. To move beyond rule-based training that repeats numerous cues and rules, we frame training as an exercise designed to elevate the degree to which individuals attend to the context in which messages are received. The training approach encourages *mindfulness* during message evaluation [46, 47]. Mindfulness is characterized by receptive attention to current surroundings and experiences [9]. Mindfulness is conceptualized as possessing state and trait components [49] and those who demonstrate higher levels of mindfulness also exhibit stronger behavioral control and self-regulation [9, 50]. Promoting mindfulness through training has served as the foundation for clinical interventions for reducing stress [25, 73], depression [80], and other psychological problems [7]. Mindfulness has also been defined and contextualized in information systems (IS) literature [10, 69]. To date, IS research investigating mindfulness has focused on its role in adoption and

postadoption behaviors (e.g., [76, 81]). In a similar vein, we argue that mindfulness can play an important role in training individuals to avoid phishing attacks. Specifically, we argue that training using mindfulness concepts will improve phishing resistance by (1) assisting individuals to dynamically allocate attention during evaluation based on whether or not a message contains an explicit request for action, (2) increasing individuals' awareness of the context through active questioning, and (3) forestalling action concerning a suspicious message. We elaborate the ways in which mindfulness concepts can improve phishing resistance below.

First, to steal information or deliver malware, phishing attack messages must encourage compliance with an explicit request for action (e.g., click on a link, download an attachment). A message that does not contain such a request requires less suspicion. Therefore, explicit requests for action represent a trigger for more mindful consideration of the context in which the message was received. In contrast to the rule-based approach, which focuses on judging the *characteristics* of a message (e.g., sender's e-mail address, makeup of an embedded link), an evaluative focus on explicit requests assists individuals in understanding the *purpose* or *intended outcome* of the message. An evaluative focus on requests in messages will likely generalize across more forms of phishing attacks and will assist individuals in allocating attention to those messages that represent the greatest risk. Informational or notification messages that do not contain an explicit request represent a much smaller risk and thus should warrant less attention.

Second, mindfulness interventions often expressly encourage individuals to pause and reflect on the context and their environment [46, 47]. Such reflection induces greater receptive attention to surroundings, which assists individuals to escape mindless processing of incoming messages, to consider each message, and to reflect on how it relates to the current context. When individuals pause to consider messages and their context, they are more likely to identify critical details such as the relationship between themselves and a message sender, relevancy of topics in a message, or the reasonableness of a request, enabling them to identify and distinguish phishing messages from legitimate messages more effectively. Reflection requires cognitive resources and time, and therefore is advisable only when a message contains an explicit request for action (e.g., click on a link, download an attachment). Reflection on the environment and the requested action may be encouraged through active questioning in which individuals pose questions to themselves about their context and the implications of their actions. Such questions direct individuals' attention to their surroundings and increase mindfulness. Langer suggests, "If you ask questions that encourage mindfulness, you bring people to the present and you're more likely to avoid an accident" [45, p. 72].

Finally, to avoid improperly jumping to conclusions, mindfulness techniques encourage forestalling judgment [8]. Forestalling judgment allows individuals to observe and take advantage of opportunities as they emerge in the environment [58]. When applied to anti-phishing training, forestalling judgment takes on new meaning. Rather than taking advantage of opportunities, forestalling judgment concerns avoiding premature labeling of suspicious messages as legitimate. When encountering

deception in general, individuals are often suspicious of deceptive messages, but are hesitant to label the messages as deception for fear of the possibility that they could be legitimate [87]. Forestalling judgment by attending to, and not suppressing suspicion, should reduce mindless acceptance of suspect messages. Additionally, when individuals are unsure about the legitimacy of a message, acknowledging suspicion could also be an impetus for verifying the suspicion with a trusted third party (e.g., IT department). Forestalling judgment and verifying suspicion may reduce individuals' reluctance to examine messages that seem suspicious or out of place but do not rise to the level of a blatant phishing attack.

Hypothesis 1: Supplemental training using mindfulness concepts will decrease the likelihood that recipients will respond to phishing attacks relative to no additional training.

Researchers and practitioners who employ the rule-based approach to anti-phishing training focus on cues that distinguish phishing messages from legitimate messages. Consistent with signal detection (e.g., [24]), training emphasized cues that were the most diagnostic (e.g., a request for confidential information, an unknown sender, and a suspicious link) and as they were identified, protective rules were developed regarding the application of these cues. These cues and rules were (and still are) widely distributed to individuals and organizations (e.g., see endnote 1). This approach to training forms an important foundation of knowledge that provides guidance to individuals about how to separate malicious from authentic messages and has been shown to reduce vulnerability to phishing attacks [14, 44].

However, the rule-based anti-phishing approach has several limitations. Rule-based training promotes behaviors that are meant to be incorporated into work habits. Organizations often repeat these rules (e.g., through annual or semiannual reminders, or mandated anti-phishing training) with the goal of strengthening protective behaviors. But, as individuals are exposed to regular reminders or repeated training, the marginal improvement conferred by the training will likely decrease as individuals become accustomed to the guidance and develop (or mistakenly believe they develop) increasing mastery over the training concepts [22]. For example, an employee who receives repeated anti-phishing reminders may not attend to and apply the requested behaviors because the employee thinks he or she already knows them. Such desensitization has been previously noted in persuasive message processing [11], and we expect it to also exist in anti-phishing training.

When rule-based training is successful in entrenching behaviors in habit, prescribed behaviors should provide consistent protection from *existing, known* phishing attacks as individuals work through their mental checklists of cues and rules every time they receive a message. However, applying consistent rules and watching for static cues assumes that the cues signaling a phishing message are stable and do not change. Though some cues may consistently indicate phishing is present (e.g., a request for action such as clicking on a link or downloading an attachment), other cues may lose their predictive power as phishing attacks become more dynamic and

customized to the targeted individuals and their activities. For example, attacks vary by the source and relevance of the source that phishers imitate [31, 82], the persuasion tactics used in the message [91], and the appearance of the message and presence of credibility cues [85]. Which sources are relevant, which tactics are effective, and which messages look convincing can change based on context and with the passage of time. Thus, individuals familiar with rule-based training may be prepared for attacks demanding a password reset from a spoofed financial institution, but may not be prepared for a drive-by attack mimicking a friend request through social media. In fact, for phishers to succeed, it is imperative that they actively avoid detection by the rule-based approach. As rules protecting individuals from phishing attacks become well-known, successful phishers seek to improve their tactics by designing messages that circumvent those rules.

Using training to elevate mindfulness during message evaluation should increase individuals' resistance to phishing attacks beyond the level achieved with familiar rule-based training. First, training encouraging mindfulness adds new instructions that will improve individuals' ability to allocate attention and attend to context during message evaluation. These skills build on the foundations delivered in previous rule-based training and help individuals perform message evaluation more efficiently and accurately. Further, if individuals are unsure about the appropriate response to a suspicious message, they are encouraged to check with a knowledgeable source rather than suppress their suspicion. Second, because mindfulness promotes the evaluation of messages in the context in which they are received and attention to any felt suspicion, the problems with strict application of mental checklists that the rule-based method instills are relaxed. As phishers develop more customized messages that do not exhibit many of the cues in rule-based training, individuals should still be able to detect the phishing message if they stop mindless processing when they encounter an explicit request and examine the alignment between message and message context.

Hypothesis 2: Supplemental training using mindfulness concepts will decrease the likelihood that recipients will respond to phishing attacks more than additional rule-based training.

Prior research investigating phishing has suggested that other characteristics of training (besides theoretical underpinning) can contribute to improved resistance to phishing attacks. Specifically, format of the training has been shown in previous research to affect phishing identification performance [44, 75]. Training presented through graphics and text has produced better learning outcomes and greater resistance to phishing attacks than training using text alone [44, 75]. Graphical presentation of information is thought to enhance information acquisition and improve performance in complex tasks [34, 35, 74]. In anti-phishing training, information acquisition and retention is especially important because detection of phishing messages is often an ancillary task that assumes a lower priority than other operationally relevant tasks. Researchers have argued that improvement from training including graphics occurs for two reasons. First, there is a closer fit between the way in which the information is presented and the mental models

in place to interpret and act on the information [16, 84]. For example, relationships between concepts (e.g., negative consequences of clicking on a link in phishing message) might be more effectively conveyed graphically than through text alone. As the task and associated information increase in abstraction, as in the case when considering suspicious messages, the fit between the format of the information and mental models may be even more important to acquisition [74]. Second, instruction using text plus graphics is thought to be more memorable and engaging and produces greater motivation to learn [53]. For example, an image of a phisher attempting to steal valuable information may be more memorable than a textual description.

Training incorporating text plus graphics has improved resistance to phishing attacks [44, 75]. However, these gains have only been demonstrated for rule-based training approaches. Therefore, we attempt to replicate these findings for both rule-based *and* mindfulness training approaches. In so doing, we hope to determine the optimal method for delivering mindfulness training and we expect that individuals exposed to training using text plus graphics should be better able to absorb and apply the training concepts and thus more accurately detect phishing messages than those exposed to text-only training.

Hypothesis 3: Training presented in a text-plus-graphics format will decrease the likelihood that participants will respond to phishing attacks more than training presented in a text-only format.

Method

Overview

To test our hypotheses, we conducted a field experiment at a university located in the Midwestern United States. We created training programs using rule-based and mindfulness approaches and randomly assigned faculty, staff, and students to participate in the training or a control task (survey-only condition without training). To test training effectiveness, we then conducted a phishing attack during which participants were directed to a fictitious website to enter their university usernames and passwords.² Data were gathered by two primary means: (1) an online survey accompanying the training and (2) participants' responses to the phishing attack.³ Our experiment followed guidelines for conducting ethical experimental phishing research [20], and we worked closely with the university administration (president's office, legal counsel), IT department (chief information officer, support staff), and the institutional review board to minimize risk to the institution and participants and ensure that relevant laws were followed.

Participants

A total of 1,048 faculty, staff, and students were invited to participate in the experiment. Participants were randomly assigned to experimental condition and were stratified such

that roughly equal numbers of faculty, staff, and students were in each condition. Participants who completed the experiment were entered into a drawing for an iPad or one of five \$50 gift cards. A total of 371 participants responded to the invitation; however, 16 of the participants did not complete the experiment and were excluded from analysis. The remaining 355 participants made up the experiment sample (a 33.9 percent completion rate). A description of the participants is provided in [Table 1](#).

The participants were familiar with phishing, with over 80 percent reporting being regularly targeted by attack messages. Participants also reported that they thought they knew what a phishing message looked like ($M = 3.8$, $SD = 1.3$, see [Table 1](#)). Additionally, conversations with university IT staff indicated that phishing attacks on the university regularly compromised accounts. The compromised accounts were used for spam or to launch more phishing attacks, which occasionally resulted in data breaches and university e-mail servers being blacklisted. As a result, the IT department regularly sent warning messages to individuals notifying them of widespread attacks. The warning messages also contained general, rule-based training on how to recognize and avoid phishing attacks. The warning messages containing rule-based training were sent at least once a semester.

Table 1. Participant Description

Phishing Familiarity (Five-Point Likert-Type Questions)	Mean	Median	S.D.
I know what a phishing message looks like.	3.8	4	1.3
I am confident in my ability to identify phishing messages.	3.7	4	1.3
I know what to do with phishing messages when I detect them.	3.7	4	1.4
Phishing Familiarity (Yes/No Questions)			
Knew someone who fell for an attack.	54%		
Never personally fallen for an attack.	70%		
Came close to falling for an attack.	16%		
Personally fell for an attack.	3%		
Unsure if personally fallen for an attack.	11%		
Self-Reported Estimate of Phishing Messages Received Per Week			
None	19%		
1–5	59%		
6–15	15%		
>16	7%		
Status		Age	
Faculty	84 (24%)	<20	26%
Staff	73 (21%)	21-29	29%
Students	198 (56%)	30-39	15%
		40-49	13%
		50-59	12%
		>60	5%

Notes: S.D. = standard deviation of the sample. Five-point Likert-type anchors (Strongly disagree to Strongly agree).

Stimulus Materials

Training Approach

Both the rule-based and mindfulness training interventions were patterned after prior training research [12, 79]. They (1) described the phenomenon, (2) described appropriate behavior in response to the phenomenon, (3) provided an opportunity to practice the behavior, and (4) provided feedback following the practice. Participants were welcomed to the training by a message from the university chief information officer, which stressed the importance of the training and thanked them for their participation. Next, participants were told that the university was experiencing an increase in phishing attacks and were shown information about the adverse consequences of responding to a phishing attack (e.g., endangering individual and organizational resources). Participants then received anti-phishing training that used either the rule-based or mindfulness approach. After training, participants were asked to practice by evaluating four e-mails (two were legitimate, two were phishing). Correct answers and explanations consistent with the training approach (rule-based or mindfulness) were provided immediately after participants submitted their answers. Training concluded with participants completing a four-item knowledge test. Again, correct answers and explanations consistent with each training approach were shown after the test was complete.

The rule-based training content was derived from anti-phishing guidance found in academic, governmental, nonprofit, and corporate sources. We first gathered rule-based guidance for individuals as set forth by previous anti-phishing researchers [44]. We then reconciled this list with resources provided by the U.S. Federal Trade Commission,⁴ the nonprofit group antiphishing.org,⁵ and anti-phishing guidelines suggested by an international bank.⁶ We identified six recommendations provided by these sources. To verify the relevance of the guidance, we presented the recommendations to two independent IT security managers employed at two different universities. They confirmed the guidance to be highly relevant in a university environment and, if followed, would substantially reduce susceptibility to phishing attacks. Therefore, these six recommendations formed the content for the rule-based training (see Table 2).

The mindfulness training was developed based on clinical mindfulness research [47] that focuses on teaching individuals to be aware of their actions, the environment, and their actions' potential consequences within the environment [48]. In past mindfulness research (e.g., [9]), reflecting before taking action was shown to predict self-regulation. Therefore, mindfulness training cautioned individuals against quickly responding to e-mail requests and encouraged them to stop, consider what e-mails ask them to do, and then take appropriate action. Consistent with the hypothesized reasoning for mindfulness training's effectiveness, it was designed around three key steps: (1) Stop! (2) Think ... (3) Check.

The first step—Stop!—was intended to teach individuals to pause any time an e-mail contains an explicit request for action (e.g., download an attachment, click on

Table 2. Recommendations for Avoiding Phishing Attacks Using the Rule-Based Approach

Recommendations
1. Never click on a link or open an attachment in an e-mail from an unknown sender.
2. Access a website by typing the web address yourself.
3. Do not reply to e-mails asking for private information.
4. Real organizations such as banks or employers will never ask for private information in an e-mail.
5. Be suspicious of a website that asks for private information.
6. Look for cues such as HTTPS in the address bar or a lock icon in your browser to identify a fake website.

a link). These requests come with some risk and the Stop! step encourages individuals to pause in order to examine the intended outcome of the message, understand potential consequences, and avoid routinized replies. The second step, Think . . ., encourages individuals to reflect on the actions they are requested to perform, the context in which the request was received, and the possible motivation for the request. Participants were instructed to consider four questions (see Table 3). By design, these questions were short and easy to remember, but they also directed attention to participants' surroundings. Finally, individuals were instructed to check. If, during the process of evaluation, any suspicion was raised, individuals were instructed to check with a trusted third party. In this training program, the local IT help desk was provided as a point of contact that would assist participants in distinguishing legitimate e-mail requests from phishing attacks. The contents and screenshots for the rule-based and mindfulness training approaches are presented in Online Appendix B.

Participants using the rule-based approach performed well on the phishing identification practice and knowledge test with average scores of 3.27 ($SD = .70$) and 3.36 ($SD = .80$), respectively. They also reported that the training was very helpful ($M = 4.10$; $SD = .85$).⁷ Likewise, participants using the mindfulness approach did well on the phishing identification practice and knowledge test with average scores

Table 3. Recommendations for Avoiding Phishing Attacks Using the Mindfulness Approach

Recommendations
1. Stop!
2. Think . . .
a. Does the request ask for private or proprietary information?
b. Is the request unexpected or rushed?
c. Does the request make sense?
d. Why would the sender need me to do this?
3. Check.

of 3.18 ($SD = .81$) and 3.53 ($SD = .82$). They also reported the training was very helpful ($M = 4.18$; $SD = .76$). Independent sample t -tests revealed no differences between participants' training test results.

Training Format

The training delivery method was manipulated such that some participants received training in a text-only format while others received training through a text-plus-graphics format. With the text-plus-graphics format, we followed the approach taken by other phishing researchers (e.g., [42, 44]) and developed a four-panel, comic-strip-like format with three characters (an e-mail user, a mentor, and a phisher) demonstrating how phishing works and how one can avoid phishing attacks. The instructional portion that presented the rule-based or mindfulness training appeared in the second panel. The text-only format contained the same content, but only the text was provided to the participants. Training materials for the text-plus-graphics and text-only formats are presented in Online Appendix B.

Phishing Attack

In coordination with the university IT department, we created a legitimate employee e-mail account for a fictitious person. We then spoofed the legitimate account when sending the phishing e-mails, which is a common practice among actual phishers [55]. We created the account for several reasons: first, we wanted to re-create as closely as possible an actual phishing attack. Second, by creating an e-mail account, we avoided using an actual employee account and creating a backlash from the phishing message against an actual university employee. Finally, we wanted control over the account to monitor correspondence (i.e., verification attempts) from experiment participants. A total of 19 individuals attempted to contact the fictitious employee to inquire about the phishing e-mail and we did not respond to the inquiries. Four individuals made repeated inquiries.

We used customizable mass e-mailing software to manage the e-mail distribution and tracked the number of e-mails that were returned as undeliverable or that bounced because of participants' e-mail client settings. The mass e-mailing software spoofed the e-mail address of the fictitious employee and sent the phishing e-mail to a randomly selected block of university students, staff, and faculty addresses every 10 minutes. No e-mails bounced or were returned as undeliverable.

Two different attack messages were used in the experiment: generic and customized. The generic message invited participants to log in to and try out a new web portal that supposedly supplied several services to university students and employees. As was done in previous phishing experiments conducted in higher education environments [90, 92], the customized message included a similar invitation, but also listed the name of the university several times (sometimes in an abbreviation), the university mascot, displayed a local phone number, and was customized to each

recipient based on their role at the university (e.g., student, staff, or faculty). Both e-mails contained a URL to the phishing website in plain text and a URL that contained a tracking number (e.g., myUniversity.org?p1234) that allowed us to track visits to the phishing website in addition to logins. No attempt was made to conceal the URL because it was a top-level .org domain that contained the acronym for the university. The generic and customized attack messages are shown in Appendix C.

Participants who followed the URL were directed to a fictitious website (shown in Appendix C) that was designed to mimic legitimate university websites. When participants typed their usernames and passwords and clicked the login button, the password was immediately deleted and was not transmitted or stored by the website. Only usernames were transmitted and recorded by the website. After participants logged in to the fictitious website, they were informed that website content was still being populated and that they should log in at a later date to view the content.

Before beginning the training, participants reported their level of expertise in identifying phishing attacks (see Table 1) and were asked for their university usernames so they would be entered into a drawing for prizes. The participants then completed the training and a questionnaire that captured experiment control variables as well as their attitudes about the training's effectiveness. Ten days following the training, participants received attack messages. The delay of 10 days between training and the phishing attack was selected as an effective balance so that participants would still recall the training, but they would not be overly sensitized to phishing attacks.

It is possible that completing the survey would prime participants to be alert to phishing e-mails, and that any reduction in phishing responses might be partially attributable to the survey rather than the training. Therefore, participants in one condition only filled out the control variable survey and did not actually receive the training. This design allowed us to isolate the effect of the training, independent of any priming effects from the survey. The no-training participants were told that the IT department was interested in attitudes about phishing and that the survey would help them understand current attitudes at the university. Table 4 presents the total number of participants in each condition.

Prior to sending the attack messages, we enlisted the help of the IT help-desk employees. We provided them with details of the experiment and gave them a script that they should use when responding to inquiries from participants. The script

Table 4. Assignment of Participants to Experimental Conditions

Conditions		Generic Phishing Message	Custom Phishing Message
No Training		33	40
Rule-based	Text	34	40
	Graphics	33	42
Mindfulness	Text	40	32
	Graphics	27	34

instructed participants not to respond to the suspicious message, to forward the message to IT help-desk employees, and to wait while the IT help-desk investigated. We did this to mitigate the risk of participants receiving actual phishing messages during the experiment and to track the number of inquiries participants made. The IT help-desk received five inquiries from participants.

Participant usernames were collected for approximately four days before an astute staff member notified the entire campus via e-mail of the phishing attack. After the phishing attack was uncovered, we sent an e-mail to all participants disclosing our involvement, the purpose of the research, and protections that safeguarded their privacy. We also shared the training programs with all university members. Usernames collected during training were matched with usernames collected by the phishing website by a researcher unaffiliated with the university. The data were then anonymized and all links to participant identities were destroyed.

Control Variables

To isolate the effects of rule-based and mindfulness training approaches, we included several control variables. First, because mindfulness has been theorized to possess both state and trait components [49], we assessed participants' e-mail mindfulness in general. This *e-mail mindfulness* scale was developed in accordance with Langer's conceptualization of mindfulness [46, 47] and other mindfulness scales in IS research [76, 81], but was specifically oriented to e-mail usage. E-mail mindfulness is a second-order factor with subcomponents: (1) alertness to distinctions, (2) orientation in the present, (3) awareness of multiple perspectives, and (4) openness to novelty. Alertness to distinction is the degree to which individuals identify differences between their e-mail practices and others' practices. Orientation to the present refers to how individuals understand the context surrounding e-mail usage or the "big picture." Awareness of multiple perspectives refers to the ability of an individual to scan an environment and identify many points of view about proper e-mail usage. Openness to novelty refers to an individual's willingness to explore new features across various situations. We also captured propensity to trust [61], perceived Internet risk [36, 52], computer self-efficacy [13], and self-reported expertise in identifying phishing messages as these variables have been examined in recent phishing research (e.g., [92]) and play an important role in training efficacy (e.g., [93]). Finally, we controlled for status (i.e., faculty, staff, student) in the university by creating two dummy variables: faculty status and student status.

A measurement model was estimated to evaluate the control variables' reliabilities as well as convergent and discriminant validities. Consistent with the conceptualization of e-mail mindfulness as a second-order factor, we followed recommendations provided by Polites et al. [65] in generating the factor scores. We used the process outlined by Wright et al. [89] to construct a second-order confirmatory factor analysis using Mplus 6.1. A reliability analysis was then performed using Cronbach's alpha and composite reliability scores [23, 88]. All alpha scores were above .70 and all composite reliability scores were above .70 [27].

All factor loadings were above the recommended .70 [27, 72]. Convergent validity was examined using the average variance extracted (AVE) for each factor. All AVEs were above the recommended .50 [21]. To establish discriminant validity, the square-root AVE values for each factor were compared with the correlations with each of the other factors and consistent with previous recommendations, the square-root AVE was greater in every case [72]. These results provide evidence of convergent and divergent validity. Appendix A displays the items and a detailed summary of measurement model tests.

Results

Forty-seven participants (13.2 percent, including those in the no-training condition) logged into the fictitious website. Descriptive statistics showing the number of participants who fell for the phishing attack are shown by condition in Table 5. Figure 1 displays the timing of the responses to the phishing attack during the first four 24 hour periods after the first e-mail was sent (10:00 a.m.). The attack generated the greatest response on the first day and decreased sharply on subsequent days. During the first 36 hours, responses were greatest in the evening.

Logistic regression was used to examine the effects of explanatory variables on whether or not participants responded to the phishing attack message. However, due to the completion rate of those invited to participate in the training and the number of control variables considered, the size of the sample was less than the recommendation for power [32]. We sacrificed power and included control variables for three reasons; first, including control variables at the expense of power is a conservative analysis approach [37]. Insignificant results may suffer from the lack of observations to detect differences; however, any significant results offer strong support for our hypotheses. Second, theory mandates the presence of some control variables (e.g., e-mail mindfulness) in the analysis. Appendix D reports fit comparisons between a model including only the control variables, a fully specified model, and the hypothesized model. The hypothesized model provides the best fit of the three models. Third, by including control variables, we explicitly rule out plausible alternative explanations for our findings. In Appendix D, we also report the analysis without control variables. This analysis satisfies sample size recommendations, and the findings remain unchanged. Table 6 presents the results of the logistic regression.

Table 5. Participant Responses to Phishing E-mail

Conditions		Generic Phishing Message ^a	Custom Phishing Message ^a
No Training		8 (24.2%)	9 (22.5%)
Rule-based	Text	4 (11.8%)	6 (15.0%)
	Graphics	5 (15.2%)	5 (11.9%)
Mindfulness	Text	3 (7.5%)	4 (12.5%)
	Graphics	2 (7.4%)	1 (2.9%)

Notes: ^aThe number of responses is shown in absolute numbers and percentages for each condition.

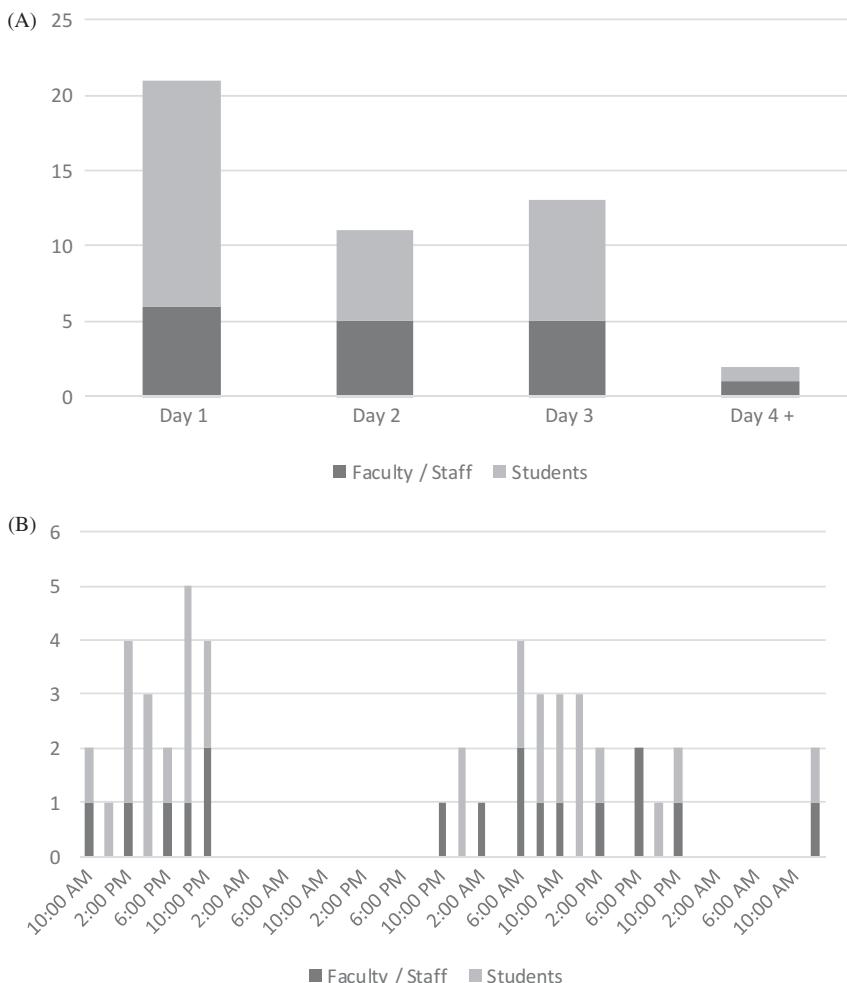


Figure 1. (a) Responses to the Phishing Attack During the First Four Days; (b) Responses to the Phishing Attack by Hours During the First Four Days.

Text and graphic training formats were equally effective in decreasing participants' likelihood to respond to phishing messages (H3: not supported). However, the mindfulness approach to training significantly reduced the likelihood that participants responded to the phishing attack (H1: supported). This effect persisted when controlling for e-mail mindfulness and other control variables. Participants who reported high e-mail mindfulness and high levels of perceptions of Internet risk were less likely to respond to the phishing attack. Also, members of the faculty were less likely to respond to the phishing attack. Students supplied their credentials 30 times (15.2 percent), staff 11 times (15.0 percent), and faculty 6 times (7.1 percent).

In a logistic regression model, explanatory variables do not have consistent, linear effects on the predicted outcome of the response variable; rather, the effect of an

Table 6. Test of Model Effects

Model factors	B (SE)	Wald	Sig.
Intercept	-1.133 (.545)	4.33	.037*
Student Status	-.121 (.490)	.06	.806
Faculty Status	-1.45 (.693)	4.41	.036*
E-mail Mindfulness	-5.14 (.959)	28.68	<.001***
Propensity to Trust	-.305 (.246)	1.53	.216
Perceived Internet Risk	-.670 (.276)	5.89	.015*
Computer Self-Efficacy	.378 (.306)	1.53	.216
Phishing Identification Expertise	-.103 (.170)	.37	.545
Phishing Customization	-.664 (.393)	2.85	.092
Rule-based Training Approach	-.419 (.513)	.67	.415
Mindfulness Training Approach	-1.51 (.568)	7.06	.008**
Graphical Training Format	-.519 (.464)	1.25	.263
Model Summary			
Omnibus test of model coefficients		<.001	
-2 log likelihood		191.76	
Cox and Snell R^2		.215	
Nagelkerke R^2		.396	
Hosmer and Lemeshow Test		.069	
Model Fit Criteria			
AIC		213.76	

Notes: * $p < .05$; ** $p < .01$; *** $p < .001$.

explanatory variable depends on the values of all the other explanatory variables in the logistic function [30]. Therefore, in Table 7, we interpret the significant coefficient for mindfulness training in the logistic regression by the magnitude of its marginal effects rather than odds ratios (see also [39, 54, 59]). In other words, to illustrate the effect of mindfulness training, its effect is estimated for experimental conditions where mindfulness training was not present while holding the other control variables constant [63]. For example, Table 7 estimates the effect of introducing mindfulness training to participants in the no-training and rule-based conditions. As significant control variables impact the marginal effect of the training approach, we display the marginal effect at the mean value and plus or minus one standard deviation from the mean for e-mail mindfulness and perceptions of Internet risks. We also display the marginal impact of the mindfulness training for faculty and nonfaculty.

To test H2, a logistic regression analysis was performed again, but participants in the no-training condition were excluded from the analysis. We then created a new variable called training approach, which had two values (rule-based approach = 0 and mindfulness approach = 1) and permitted a direct comparison between the two approaches. A significant, negative coefficient would indicate that the mindfulness

Table 7. Marginal Effect of Introducing Mindfulness Training

Control variable values	Conditions	Probability of responding	Marginal effect of mindfulness training ^a
E-mail Mindfulness^b			
Low E-mail Mindfulness (mean - 1 StDev = -.33)	No training	.444	-.294
	Rule-based	Text	.345
		Graphics	.238
	Mindfulness	Text	.150
		Graphics	.095
Mean E-mail Mindfulness (mean = .00)	No training	.128	-.097
	Rule-based	Text	.088
		Graphics	.054
	Mindfulness	Text	.031
		Graphics	.019
High E-mail Mindfulness (mean + 1 StDev = .33)	No training	.026	-.020
	Rule-based	Text	.017
		Graphics	.010
	Mindfulness	Text	.006
		Graphics	.004
Perceptions of Internet Risk^c			
Low Internet Risk (mean - 1 StDev = -.75)	No training	.430	-.287
	Rule-based	Text	.331
		Graphics	.228
	Mindfulness	Text	.143
		Graphics	.090
Mean Internet Risk (mean = .00)	No Training	.313	-.222
	Rule-based	Text	.231
		Graphics	.151
	Mindfulness	Text	.092
		Graphics	.057
High Internet Risk (mean + 1 StDev = .75)	No training	.216	-.159
	Rule-based	Text	.154
		Graphics	.097
	Mindfulness	Text	.058
		Graphics	.035
University Status^d			
Faculty	No training	.107	-.081
	Rule-based	Text	.073
		Graphics	.045
	Mindfulness	Text	.026
		Graphics	.016
Staff	No training	.340	-.238
	Rule-based	Text	.253
		Graphics	.168
	Mindfulness	Text	.102
		Graphics	.063

(continues)

Table 7. Continued

Control variable values	Conditions	Probability of responding	Marginal effect of mindfulness training ^a
Students	No training	.313	-.222
	Rule-based Text	.231	-.169
	Rule-based Graphics	.151	-.114
	Mindfulness Text	.092	
	Mindfulness Graphics	.057	

Notes: ^aMarginal effects are shown in terms of probabilities.

^b In calculating the marginal impact of mindfulness training, phishing customization and a student population are assumed. Mean values for propensity to trust, perceived Internet risk, computer self-efficacy, and phishing identification expertise are also assumed.

^c In calculating the marginal impact of mindfulness training, phishing customization and a student population are assumed. Mean values for e-mail mindfulness, propensity to trust, computer self-efficacy, and phishing identification expertise are also assumed.

^d In calculating the marginal impact of mindfulness training, phishing customization is assumed. Mean values for e-mail mindfulness, propensity to trust, perceived Internet risk, computer self-efficacy, and phishing identification expertise are also assumed.

Table 8. Test of Model Effects Comparing Mindfulness to Rule-Based Approaches

Model Factors	B (SE)	Wald	Sig.
Intercept	-1.204 (.555)	4.70	.030
Student Status	-.442 (.545)	.66	.417
Faculty Status	-2.321 (1.124)	4.27	.039*
E-mail Mindfulness	-3.497 (.990)	12.46	<.001***
Propensity to Trust	-.215 (.272)	.62	.429
Perception of Internet Risk	-.325 (.308)	1.31	.253
Computer Self-Efficacy	.356 (.345)	1.06	.303
Phishing Identification Expertise	-.238 (.194)	1.51	.219
Phishing Customization	-.400 (.441)	.82	.365
Training Approach	-.978 (.467)	4.39	.036*
Graphical Training Format	-.417 (.446)	.88	.349
Model Summary			
Omnibus test of model coefficients		<.001	
-2 log likelihood		149.05	
Cox and Snell R^2		.139	
Nagelkerke R^2		.282	
Hosmer and Lemeshow test		.871	
Model Fit Criteria			
AIC		207.77	

Notes: * $p < .05$; ** $p < .01$; *** $p < .001$.

approach reduced phishing response above the effect of the rule-based approach. As Table 8 illustrates, our finding supports H2.

Discussion

This research advances understanding of how individuals can be trained to resist phishing attacks. Consistent with prior research (e.g., [3, 43]), we demonstrated that training reduced individuals' susceptibility to phishing attacks. However, individuals' past training experience and the theoretical approach underpinning the training significantly influenced how effective the additional training was. We outline important implications of our results for theory and practice.

Implications for Theory

First, the results support our assertion that once individuals are familiar with phishing and the cues and rules they might use to avoid attacks, additional rule-based training may be less effective than other training approaches. This decline can spring from several sources such as desensitization, mistaken belief in existing expertise, or forgetting, and warrants additional attention from researchers to understand how and when this decline occurs. This finding also highlights an important paradox for organizations as leaders consider ways to mitigate phishing risks through repeated training: individuals may feel prepared to protect against phishing attacks, but their actual behavior suggests that they may not be. Participants in the no-training and rule-based training conditions believed that they were well-equipped to counter phishing attacks, reporting that they knew what a phishing message looked like ($M = 3.55$; $SD = 1.30$) and expressing confidence in their ability to detect them ($M = 3.54$; $SD = 1.28$). We note that had participants followed the guidance provided in the additional rule-based training (e.g., never click on a link or open an attachment in an e-mail from an unknown sender; be suspicious of a website that asks for private information), they would have avoided both the customized and generic phishing attacks. Nevertheless, a total of 20 participants (13.4 percent) in the rule-based training condition responded to the phishing attack. Participants believed they knew how to avoid phishing messages, but an examination of their behavior revealed that not all of them did and the security of the entire organization could have been imperiled as a result.

Participants in the mindfulness condition also reported a high level of expertise ($M = 3.52$; $SD = 1.31$) and confidence ($M = 3.47$; $SD = 1.29$) in identifying phishing messages. In contrast to the results from the rule-based training condition, participants with supplemental mindfulness training demonstrated lower susceptibility to phishing attacks (10 participant responses (7.5 percent)). These results suggest that supplemental mindfulness training may reach individuals who already think they can identify phishing messages, but may not be doing so. The results also suggest that training may affect attitudes and behavior differently. For example, training may

sustain confidence and perceived expertise, but it may not produce a corresponding increase in protective behaviors. The results underscore the need for research examining both attitudes and actual behavior in IT security research.

Second, in referent literature outside of IT security, there has been debate about the efficacy of mindfulness training (e.g., [7]). The techniques included in the anti-phishing training were designed to promote mindfulness during message evaluation: (1) pause and focus on the actions being requested; (2) observe the environment (perhaps with the direction of questions intended to promote deeper consideration of the environment); (3) forestall judgment if suspicion is raised and take action to verify the suspicion. Our findings join a growing body of evidence (e.g., [18, 71]) suggesting that mindfulness techniques can be successfully taught to individuals and that the results of the training rise above mere awareness to the level of behavior. Both the rule-based condition, which was based on contemporary anti-phishing research and the no-training condition induced awareness of phishing. Yet only the supplemental mindfulness training yielded a significant reduction in responses to the phishing attacks. The three-step mindfulness technique may form the foundation of other training interventions both in IT security contexts and in alternative application areas in which individuals must manage several competing priorities for attention. Our findings also provide evidence that successful training in mindfulness techniques can be digitally delivered. Recent cyber-security awareness campaigns are now encouraging individuals to pause and consider the risks of their actions.⁸ Our findings support this direction and open the door to other mindfulness-related interventions that could be digitally delivered.

Third, we have provided evidence of the efficacy of mindfulness theory as a useful lens through which researchers can draw insight into how to further improve training and resistance to phishing attacks. Phishing has been traditionally treated as a detection problem, with corresponding training focusing on cues and rules that support detection. Mindfulness theory represents a new category of theories germane to anti-phishing training. Its focus on the allocation of cognitive resources enables the development of strategies and mental models that help individuals detect phishing messages while balancing the demands of organizational responsibilities. Our findings suggest that once individuals have been exposed to rules for separating phishing from legitimate messages, principles of mindfulness may effectively be layered on to further increase resistance to phishing attacks. Although we chose mindfulness theory because of the way in which its unique tenets (e.g., observing the environment, forestalling judgment) align with detecting phishing messages, we acknowledge that there are many other theories that could explain the allocation of cognitive resources and the creation of mental models (e.g., dual process theories, signaling theory). We view the success of mindfulness theory in phishing training as an example of how cognitive theories might be used to improve the effectiveness that has already been achieved by training focusing on detection. Signal detection and cognitive theories address different, but complementary areas and our results suggest that further investigation of how cognitive theories could be applied in IT security training could be fruitful.

Fourth, our findings suggest that content, not format, contributes to anti-phishing training's effectiveness. In addition to comparing two training approaches, we evaluated whether their effectiveness varied with two training formats: text-only and text-plus-graphics. Consistent with prior research (e.g., [44, 75]), we hypothesized that a text-plus-graphics format would be more memorable and more effective in teaching users how to recognize and respond to phishing. However, the text-only and text-plus-graphics formats were equally effective. A potential explanation for the equivalent effectiveness of the training formats may stem from the analyzable nature of phishing detection. Since participants had previously been exposed to anti-phishing training, they likely shared a common understanding of what was required to detect phishing messages. Prior research has found that when a common understanding of the correct way to perform a task exists (i.e., the task is analyzable), text-only and multimedia formats perform equally well [40, 51]. Such common understanding may not have been as widespread in earlier phishing research.

Finally, our findings illuminate the nomological network that shapes responses to phishing. Our results show that the majority of respondents were not enticed by the phishing messages, similar to results of prior phishing experiments [56]. Furthermore, our analysis of control variables suggests that a population can be segmented to identify individuals who are most likely to respond to a phishing attack. Not surprisingly, participants who were low in e-mail mindfulness and who perceived a low level of risk from performing tasks on the Internet were highly susceptible to phishing attacks. Status at the university also influenced the probability that an individual would respond to a phishing message. These findings are consistent with past research indicating that age and education assist individuals in avoiding phishing attacks [31]. Our findings also suggest that training using mindfulness techniques dramatically reduces the susceptibility of these vulnerable groups. For example, as shown in [Table 7](#), for students who exhibit low e-mail mindfulness, training using mindfulness techniques can reduce the probability of responding to a phishing attack by 29.4 percent.

Implications for Practice

Our results have several implications for practice. First, this study underscores the importance of training individuals on how to detect and respond to suspected phishing e-mails. Individuals' ability to identify and avoid phishing attacks is a critical line of defense against phishing. However, while our results suggest that the addition of mindfulness training may result in fewer successful attacks, managers should be realistic in their expectations about phishing training. The use of mindfulness training, as our results demonstrate, does not drive responses to phishing attacks down to zero. Therefore, training is best positioned as part of a layered set of defenses that aligns automated tools (e.g., automated filtering and warnings), secure organizational practices (e.g., policies concerning e-mailing links and using centralized document storage), and training components.

Second, results show that it is crucial to consider the theoretical base that underpins a training program. The rule-based and mindfulness approaches frame anti-phishing training in different ways, and understanding the differences and complementarities are useful in building resistance to phishing attacks. Further, individuals will not all benefit to the same degree from mindfulness training. Individuals who have less experience and education, are low in e-mail mindfulness, and are low in perceptions of Internet risk will benefit disproportionately from the training while those who have more experience and education, are high in e-mail mindfulness, and are high in perceptions of Internet risk will not derive as much benefit. Therefore, when allocating training resources or evaluating training programs, those charged with managing organizational security should consider not only the training's theoretical base but also the demographics of their organization and how the training's theoretical base aligns with perceptions and experience of organization members.

Third, we demonstrate that anti-phishing training delivered via relatively brief, online sessions can be effective. Therefore, managers could potentially leverage our methods to create and distribute anti-phishing training electronically to a large number of users. Our results imply that the presentation of the training need not be complex or costly. Results show that training consisting of only text was just as effective as training shown via a text-plus-graphics presentation method. Additionally, given that the content for supplemental mindfulness training is relatively simple, it could easily be adapted for other presentation methods.

Fourth, the supplemental mindfulness training we have developed and tested offers significant benefits to individuals. We offer detailed explanations and illustration of our training materials in Appendix B and invite researchers and practitioners to use, modify, and expand these training materials to improve phishing mitigation efforts.

Limitations and Future Research

This research has several limitations that provide opportunities for future research. First, we present mindfulness training as a possible companion to (not a substitution for) rule-based training that most organizations have in place. We evaluated additional rule-based and supplemental mindfulness training as two distinct approaches to understanding their comparative effectiveness. Given that most organizations offer some form of rule-based training, future research could investigate how these training approaches might be integrated with each other, and perhaps with other approaches using new theory, to maintain interest and consistent application of protective behaviors.

Additionally, we did not examine how long the effect of mindfulness training lasts or the effect of mindfulness training after repeated exposure. As with rule-based training, we expect that mindfulness training may demonstrate diminishing marginal effectiveness as individuals become increasingly familiar with training content. Future research could investigate how training's diminishing marginal effectiveness appears and what measures can be taken to forestall it.

Next, we acknowledge that our sample is composed of individuals who responded to online solicitations to participate in the training. While demonstrating effectiveness of training for individuals who respond to online solicitations is a valuable improvement, the participants may not reflect the practices of the general population. Mindfulness methods have been theorized to provide reliable outcomes in other IS contexts (e.g., systems management [10]). Hence, as is common for research on training, our study needs replication in diverse contexts.

Future research should more thoroughly consider the tension between information security policies and regular organizational practices. For example, one commonly used information security rule is to not click on links or attachments from unknown senders, but daily activity in many firms likely results in the sending and receiving of many links and attachments within and between organizations in the course of regular business. Future research should consider how to craft anti-phishing policies and how to train organizational e-mail *senders* to adhere to good security policies such that poor security behavior does not become mindlessly ingrained among users.

Finally, a mindfulness approach to training may be helpful in many other contexts aside from phishing. Since mindfulness techniques encourage active awareness of context, training could easily be adapted to encourage appropriate behavior in other IT security settings (e.g., compliance with security policy, disclosure on social media).

Conclusion

Phishing is a scourge of modern organizations that is only expected to grow. Organizations have acted to sensitize individuals to the danger of phishing and equip them to be able to identify and avoid attacks. With this research, we probed the efficacy of training approaches and offer additional means to combat the threat of phishing. Mindfulness (among other cognitive theories) represents a valuable avenue by which we can build interventions that will reduce individuals' susceptibility to phishing attacks and thus enhance information security.

Acknowledgments: The authors gratefully acknowledge the assistance of three anonymous reviewers. We also acknowledge participants in research presentations at the University of Oklahoma, Colorado State University, University of British Columbia, University of Dayton, University of Georgia, HEC Montreal, and Temple University.

Supplemental File

Supplemental data for this article can be found on the publisher's website at [10.1080/07421222.2017.1334499](https://doi.org/10.1080/07421222.2017.1334499)

NOTES

1. Federal Trade Commission: <https://www.consumer.ftc.gov/articles/0003-phishing>; Microsoft: <https://www.microsoft.com/en-us/safety/online-privacy/phishing-symptoms.aspx>

Chase Bank: <https://www.chase.com/digital/resources/privacy-security/security/suspicious-emails>; Anti-Phishing Working Group: <http://apwg.org/resources/Educate-Your-Customers/> (All accessed on June 25, 2017).

2. The university username and password grant access to university e-mail and university resources (e.g., central IT resources, human resource records, academic records).
3. We piloted the instruments using student participants at a different university and found the scales in the survey valid and reliable (see [23]).
4. <https://www.consumer.ftc.gov/articles/0003-phishing> (accessed on June 25, 2017)
5. <http://apwg.org/resources/Educate-Your-Customers/> (accessed on June 25, 2017)
6. <https://www.chase.com/digital/resources/privacy-security/security/suspicious-emails> (accessed on June 25, 2017)
7. The identification practice and knowledge test were scored out of four. The item about the training helpfulness was “This training helped me learn how to identify phishing messages,” and was scored on a five-point scale with Strongly Disagree and Strongly Agree as endpoints.
8. <https://www.staysafeonline.org/stop-think-connect/> (accessed on June 25, 2017)

REFERENCES

1. Abbasi, A.; Zahedi, F.; Zeng, D.; Chen, Y.; Chen, H.; and Nunamaker, J.F. Enhancing predictive analytics for anti-phishing by exploiting website genre information. *Journal of Management Information Systems*, 31, 4 (2015), 109–157.
2. Almomani, A.; Gupta, B.; Atawneh, S.; Meulenberg, A.; and Almomani, E. A survey of phishing email filtering techniques. *IEEE Communications Surveys and Tutorials*, 15, 4 (2013), 2070–2090.
3. Alnajim, A., and Munro, M. An anti-phishing approach that uses training intervention for phishing websites detection. Paper presented at the *Sixth International Conference on Information Technology: New Generations*, 2009. *ITNG '09*, 2009, pp. 405–410.
4. Anderson, B.B.; Kirwan, C.B.; Jenkins, J.L.; Eargle, D.; Howard, S.; and Vance, A. How polymorphic warnings reduce habituation in the brain: Insights from an fMRI study. Paper presented at *CHI*, ACM, Seoul, Korea, 2015.
5. Anti-Phishing Working Group. Phishing activity trends report. In APWG (ed.), Anti-Phishing Working Group, 2016. Available at: docs.apwg.org/reports/apwg_trends_report_q1_2016.pdf (accessed on June 25, 2017)
6. Ashby, F.G.; Maddox, W.T.; and Bohil, C. Observational versus feedback training in rule-based and information-integration category learning. *Memory and Cognition*, 30, 5 (2002), 666–677.
7. Baer, R.A. Mindfulness training as a clinical intervention: A conceptual and empirical review. *Clinical Psychology: Science and Practice*, 10, 2 (2003), 125–143.
8. Baer, R.A.; Smith, G.T.; and Allen, K.B. Assessment of mindfulness by self-report: The Kentucky Inventory of Mindfulness Skills. *Assessment*, 11, 3 (2004), 191–206.
9. Brown, K.W.; Ryan, R.M.; and Creswell, J.D. Mindfulness: Theoretical foundations and evidence for its salutary effects. *Psychological Inquiry*, 18, 4 (2007), 211–237.
10. Butler, B.S., and Gray, P.H. Reliability, mindfulness, and information systems. *MIS Quarterly*, 30, 2 (2006), 211–224.
11. Cacioppo, J.T., and Petty, R.E. Effects of message repetition and position on cognitive response, recall, and persuasion. *Journal of Personality and Social Psychology*, 37, 1 (1979), 97–109.
12. Compeau, D.R., and Higgins, C.A. Application of social cognitive theory to training for computer skills. *Information Systems Research*, 6, 2 (1995), 118–143.
13. Compeau, D.R., and Higgins, C.A. Computer self-efficacy: Development of a measure and initial test. *MIS Quarterly*, 19, 2 (1995), 189–211.
14. Cranor, L.F. Can phishing be foiled? *Scientific American*, 299, 6 (2008), 104–110.

15. D'Arcy, J.; Herath, T.; and Shoss, M.K. Understanding employee responses to stressful information security requirements: A coping perspective. *Journal of Management Information Systems*, 31, 2 (2014), 285–318.
16. Dennis, A.R., and Carte, T.A. Using geographical information systems for decision making: Extending cognitive fit theory to map-based presentations. *Information Systems Research*, 9, 2 (1998), 194–203.
17. Dhamija, R.; Tygar, J.D.; and Hearst, M. Why phishing works. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. Montreal, Canada: ACM, 2006, pp. 581–590.
18. Eberth, J., and Sedlmeier, P. The effects of mindfulness meditation: A meta-analysis. *Mindfulness*, 3, 3 (2012), 174–189.
19. Federal Bureau of Investigation. FBI warns of dramatic increase in business e-mail scams. 2016. Available at: <http://fbi.gov/contact-us/field-offices/phoenix/news/press-releases/fbi-warns-of-dramatic-increase-in-business-e-mail-scams> (accessed on June 25, 2017)
20. Finn, P., and Jakobsson, M. Designing and conducting phishing experiments. *IEEE Technology and Society*, 6, 2 (2008), 66–68.
21. Fornell, C., and Larcker, D.F. Evaluating structural equations models with unobservable variables and measurement error. *Journal of Marketing Research*, 18, 1 (1981), 39–50.
22. Fuller, C.M.; Biros, D.P.; and Imperial, M.J. Knowledge retention in information assurance computer-based training: A comparative study of two courses for network user training. In *Proceedings of the Sixth Annual Security Conference*, Las Vegas, NV: Virginia Commonwealth University, 2007, pp. 28-1–28-14.
23. Gefen, D.; Straub, D.W.; and Rigdon, E.E. An update and extension to SEM Guidelines for Administrative and Social Science Research. *MIS Quarterly*, 35, 2 (2011), iii–xiv.
24. Green, D.M., and Swets, J.A. *Signal Detection Theory and Psychophysics*. New York, NY: Wiley, 1966.
25. Grossman, P.; Niemann, L.; Schmidt, S.; and Walach, H. Mindfulness-based stress reduction and health benefits: A meta-analysis. *Journal of Psychosomatic Research*, 57, 1 (2004), 35–43.
26. Guo, K.H.; Yuan, Y.; Archer, N.P.; and Connelly, C.E. Understanding nonmalicious security violations in the workplace: A composite behavior model. *Journal of Management Information Systems*, 28, 2 (2011), 203–236.
27. Hair, J.F. Jr.; Anderson, R.E.; Tatham, R.L.; and Black, W.C. *Multivariate Data Analysis with Readings*. Englewood Cliffs, NJ: Prentice Hall, 1998.
28. Harbison, C. 10 largest data breaches of 2014; The Sony hack is not one of them! iDigital Times, Available at: <http://www.idigitaltimes.com/10-largest-data-breaches-2014-sony-hack-not-one-them-403219> (accessed on June 25, 2017)
29. Heartfield, R., and Loukas, G. A Taxonomy of attacks and a survey of defence mechanisms for semantic social engineering attacks. *ACM Computing Surveys (CSUR)*, 48, 3 (2015), 37-1–37-39.
30. Hoetker, G. The use of logit and probit models in strategic management research: Critical issues. *Strategic Management Journal*, 28, 4 (2007), 331–343.
31. Hong, J. The state of phishing attacks. *Communications of the ACM*, 55, 1 (2012), 74–81.
32. Hosmer, D.W., and Lemeshow, S. *Applied Logistic Regression*. New York, NY: Wiley, 2000.
33. Jackson, C.; Simon, D.; Tan, D.; and Barth, A. An evaluation of extended validation and picture-in-picture phishing attacks. In S. Dietrich and R. Dhamija (eds.), *Financial Cryptography and Data Security*. Berlin, Germany: Springer, 2007, pp. 281–293.
34. Jarvenpaa, S.L. The effect of task demands and graphical format on information processing strategies. *Management Science*, 35, 3 (1989), 285–303.
35. Jarvenpaa, S.L., and Dickson, G.W. Graphics and managerial decision making: research based guidelines. *Communications of the ACM*, 31, 6 (1988), 764–774.
36. Jarvenpaa, S.L.; Tractinsky, N.; and Saarinen, L. Consumer trust in an Internet store: A cross-cultural validation. *Journal of Computer-Mediated Communication*, 5, 2 (1999), 1–35.

37. Jo, B. Statistical power in randomized intervention studies with noncompliance. *Psychological Methods*, 7, 2 (2002), 178–193.
38. Johnson, J. If 2014 was the year of the data breach, brace for more. *Forbes*, 2015. Available at: <https://www.forbes.com/sites/danielfisher/2015/01/02/if-2014-was-the-year-of-the-data-breach-brace-for-more> (accessed on June 25, 2017)
39. Kaufman, R.L. Comparing the effects of dichotomous logistic regression: A variety of standardized coefficients. *Social Science Quarterly*, 77, 1 (1996), 90–109.
40. Kelton, A.S.; Pennington, R.R.; and Tuttle, B.M. The effects of information presentation format on judgment and decision making: A review of the information systems research. *Journal of Information Systems*, 24, 2 (2010), 79–105.
41. Kirk, J. Ham-fisted phishing attack seeks LinkedIn logins. *CSO Magazine*, 2015. Available at: <http://www.csionline.com/article/2868889/cyber-attacks-espionage/hamfisted-phishing-attack-seeks-linkedin-logins.html> (accessed on June 25, 2017)
42. Kumaraguru, P.; Cranshaw, J.; Acquisti, A.; Cranor, L.; Hong, J.; Blair, M.A.; and Pham, T. School of phish: A real-world evaluation of anti-phishing training. In *SOUPS '09 Proceedings of the Fifth Symposium on Usable Privacy and Security*. Mountain View, CA: ACM, 2009, pp. 1–12.
43. Kumaraguru, P.; Rhee, Y.; Acquisti, A.; Cranor, L.; Hong, J.; and Nunge, E. The design and evalauation of an embedded training email systems. *Computer Human Interaction (CHI)*. San Jose, CA: ACM Press, 2007, pp. 905–914.
44. Kumaraguru, P.; Sheng, S.; Acquisti, A.; Cranor, L.F.; and Hong, J. Teaching Johnny not to fall for phish. *ACM Transactions on Internet Technology (TOIT)*, 10, 2 (2010), 7-1–7-31.
45. Langer, E. Mindfulness in the age of complexity. *Harvard Business Review*, 92, 3 (2014), 68–73.
46. Langer, E.J. *Mindfulness*. Reading, MA: Addison-Wesley, 1989.
47. Langer, E.J. *The Power of Mindful Learning*. Reading, MA: Addison-Wesley 1997.
48. Langer, E.J., and Piper, A. The Prevention of Mindlessness. *Journal of Personality and Social Psychology*, 53 (1987), 280–287.
49. Lau, M.A.; Bishop, S.R.; Segal, Z.V.; Buis, T.; Anderson, N.D.; Carlson, L.; Shapiro, S.; Carmody, J.; Abbey, S.; and Devins, G. The Toronto mindfulness scale: Development and validation. *Journal of Clinical Psychology*, 62, 12 (2006), 1445–1467.
50. Leary, M.R.; Adams, C.E.; and Tate, E.B. Hypo-egoic self-regulation: Exercising self-control by diminishing the influence of the self. *Journal of Personality*, 74, 6 (2006), 1803–1831.
51. Lim, K.H., and Benbasat, I. The effect of multimedia on perceived equivocality and perceived usefulness of information systems. *MIS Quarterly*, 24, 3 (2000), 449–471.
52. Malhotra, N.K.; Kim, S.S.; and Agarwal, J. Internet users' information privacy concerns (IUIPC): The construct, the scale, and a causal model. *Information Systems Research*, 15, 4 (2004), 336–355.
53. Mayer, R.E. *Multimedia Learning*. New York, NY: Cambridge University Press, 2001.
54. Meservy, T.O.; Jensen, M.L.; and Fadel, K. Evaluation of competing candidate solutions in electronic networks of practice. *Information Systems Research*, 25, 1 (2014), 15–34.
55. Mohammad, R.M.; Thabtah, F.; and McCluskey, L. Tutorial and critical analysis of phishing websites methods. *Computer Science Review*, August (2015), 1–24.
56. Mohebzada, J.G.; El Zarka, A.; Bhojani, A.H.; and Darwish, A. Phishing in a university community: Two large scale phishing experiments. 2012 International Conference on Innovations in Information Technology (IIT), IEEE, 2012, pp. 249–254.
57. Myers, S. Introduction to phishing. In M. Jakobsson and S. Myers (eds.), *Phishing and Countermeasures: Understanding the Increasing Problem of Electronic Identity Theft*. Hoboken, NJ: Wiley, 2007, pp. 1–29.
58. Ndubisi, N.O. Mindfulness, reliability, pre-emptive conflict handling, customer orientation and outcomes in Malaysia's healthcare sector. *Journal of Business Research*, 65, 4 (2012), 537–546.
59. Norton, E.C.; Wang, H.; and Ai, C. Computing interaction effects and standard errors in logit and probit models. *Stata Journal*, 4, 2 (2004), 154–167.

60. Orlikowski, W.J., and Yates, J. Genre repertoire: The structuring of communicative practices in organizations. *Administrative Science Quarterly*, 39, 4 (1994), 541–574.
61. Pavlou, P.A., and Gefen, D. Building effective online marketplaces with institution-based trust. *Information Systems Research*, 15, 1 (2004), 37–59.
62. Perez, E., and Prokupecz, S. How the U.S. thinks Russians hacked the White House. CNN, 2015. Available at: <http://www.cnn.com/2015/04/07/politics/how-russians-hacked-the-wh> (accessed on June 25, 2017)
63. Petersen, T. A comment on presenting results from logit and probit models. *American Sociological Review*, 50, 1 (1985), 130–131.
64. Png, I.L., and Wang, Q. Information security: Facilitating user precautions vis-à-vis enforcement against attackers. *Journal of Management Information Systems*, 26, 2 (2009), 97–121.
65. Polites, G.; Roberts, N.; and Thatcher, J. Conceptualizing models using multidimensional constructs: A conceptual review and guidelines for their use. *European Journal of Information Systems*, 21, 1 (2012), 22–48.
66. Polites, G.L., and Karahanna, E. The embeddedness of information systems habits in organizational and individual level routines: Development and disruption. *MIS Quarterly*, 37, 1 (2013), 221–246.
67. Puhakainen, P., and Siponen, M. Improving employees' compliance through information systems security training: An action research study. *MIS Quarterly*, 34, 4 (2010), 757–778.
68. Purkait, S. Phishing counter measures and their effectiveness: Literature review. *Information Management and Computer Security*, 20, 5 (2012), 382–420.
69. Roberts, N.; Thatcher, J.; and Klein, R. Tying context to post-adoption behavior with information technology: A conceptual and operational definition of mindfulness. *AMCIS 2007 Proceedings*, Keystone, CO, 2007, pp. 1–6.
70. Savvas, A. 91% of cyberattacks begin with spear phishing email. *TechWorld*. TechWorld, 2012. Available at: <http://www.techworld.com/news/security/91-of-cyberattacks-begin-with-spear-phishing-email-3413574/> (accessed on June 25, 2017)
71. Sedlmeier, P.; Eberth, J.; Schwarz, M.; Zimmermann, D.; Haarig, F.; Jaeger, S.; and Kunze, S. The psychological effects of meditation: A meta-analysis. *Psychological Bulletin*, 138, 6 (2012), 1139–1171.
72. Segars, A. Assessing the unidimensionality of measurement: A paradigm and illustration within the context of information systems research. *Omega*, 25, 1 (1997), 107–121.
73. Shapiro, S.L.; Schwartz, G.E.; and Bonner, G. Effects of mindfulness-based stress reduction on medical and premedical Students. *Journal of Behavioral Medicine*, 21, 6 (1998), 581–599.
74. Speier, C. The influence of information presentation formats on complex task decision-making performance. *International Journal of Human-Computer Studies*, 64, 11 (2006), 1115–1131.
75. Srikwan, S., and Jakobsson, M. Using cartoons to teach internet security. *Cryptologia*, 32, 2 (2008), 137–154.
76. Sun, H.; Fang, Y.; and Zhou, H. Choosing a fit technology: Understanding mindfulness in technology adoption and continuance. *Journal of the Association for Information Systems*, 17, 6 (2016), 377–412.
77. Symantec. Internet security threat report 2013. Mountain View, CA: Symantec Corporation, 2013. Available at: http://www.symantec.com/content/en/us/enterprise/other_resources/b-istr_main_report_v18_2012_21291018.en-us.pdf (accessed on June 25, 2017)
78. Symantec. Internet security threat report 2014. Vol. 189, 2014. Available at: [symantec.com/content/en/us/enterprise/other_resources/b-istr_main_report_v19_21291018.en-us.pdf](http://www.symantec.com/content/en/us/enterprise/other_resources/b-istr_main_report_v19_21291018.en-us.pdf) (accessed on June 25, 2017)
79. Taylor, P.J.; Russ-Eft, D.F.; and Chan, D.W.L. A meta-analytic review of behavior modeling training. *Journal of Applied Psychology*, 90, 4 (2005), 692–709.
80. Teasdale, J.D.; Williams, J.M.G.; Soulsby, J.M.; Segal, Z.V.; Ridgeway, V.A.; and Lau, M.A. Prevention of relapse/recurrence in major depression by mindfulness-based cognitive therapy. *Journal of Consulting and Clinical Psychology*, 68, 4 (2000), 615–623.

81. Thatcher, J.B.; Wright, R.T.; Sun, H.; Klein, R.; and Zagencyk, T. Mindfulness in information technology use: A conceptual and operational definition. *MIS Quarterly*, forthcoming.
82. van der Merwe, A.; Loock, M.; and Dabrowski, M. Characteristics and responsibilities involved in a phishing attack. In *Proceedings of the Fourth International Symposium on Information and Communication Technologies*. Cape Town, South Africa: Trinity College Dublin, 2005, pp. 249–254.
83. Vance, A.; Elie-Dit-Cosaque, C.; and Straub, D.W. Examining trust in information technology artifacts: The effects of system quality and culture. *Journal of Management Information Systems*, 24, 4 (2008), 73–100.
84. Vessey, I., and Galletta, D. Cognitive fit: An empirical study of information acquisition. *Information Systems Research*, 2, 1 (1991), 63–84.
85. Vishwanath, A. Mobile device affordance: Explicating how smartphones influence the outcome of phishing attacks. *Computers in Human Behavior*, 63 (2016), 198–207.
86. Vishwanath, A.; Herath, T.; Chen, R.; Wang, J.; and Rao, H.R. Why do people get phished? Testing individual differences in phishing vulnerability within an integrated, information processing model. *Decision Support Systems*, 51, 3 (2011), 576–586.
87. Vrij, A. Nonverbal communication and deception. In V. Manusov and M.L. Patterson (eds.), *The Sage Handbook of Nonverbal Communication*. Thousand Oaks, CA: Sage, 2006, pp. 341–359.
88. Werts, C.E.; Linn, R.L.; and Joreskog, K. Interclass reliability estimates: Testing structural assumptions. *Educational and Psychological Measurement*, 34, 1 (1974), 25–33.
89. Wright, R.T.; Campbell, D.E.; Thatcher, J.B.; and Roberts, N. Operationalizing multi-dimensional constructs in structural equation modeling: Recommendations for IS research. *Communications of the Association for Information Systems*, 40 (2012), 367–412.
90. Wright, R.T.; Chakraborty, S.; Basoglu, A.; and Maret, K. Where did they go right? Understanding the deception in phishing communications. *Group Decision and Negotiation*, 19, 4 (2010), 391–416.
91. Wright, R.T.; Jensen, M.L.; Thatcher, J.; Dinger, M.; and Maret, K. Influence techniques in phishing attacks: An examination of vulnerability and resistance. *Information Systems Research*, 25, 2 (2014), 385–400.
92. Wright, R.T., and Maret, K. The influence of experiential and dispositional factors in phishing: An empirical investigation of the deceived. *Journal of Management Information Systems*, 27, 1 (2010), 273–303.
93. Yi, M.Y., and Davis, F.D. Improving computer training effectiveness for decision technologies: Behavior modeling and retention enhancement. *Decision Sciences*, 32, 3 (2001), 521–544.