WHO BITES THE HOOK? INVESTIGATING EMPLOYEES' SUSCEPTIBILITY TO PHISHING: A RANDOMIZED FIELD EXPERIMENT

Research Paper

Franz, Anjuli, Technical University of Darmstadt, Germany, franz@ise.tu-darmstadt.de

Croitor, Evgheni, Technical University of Darmstadt, Germany, croitor@ise.tu-darmstadt.de

Abstract

Phishing is a major threat to organizational information security, with the employee being a critical link in the security chain. Understanding why employees fall for phish is therefore crucial in order to design effective countermeasures. In this article, we investigate how (1) employees' social networking site (SNS) use and (2) message involvement affects their susceptibility to phishing attacks. SNS use has been found to strongly influence users' information processing, exacerbating proneness to heuristic decision-making, and hence making them easy prey for criminals' influence techniques. We present a model to investigate the moderating role of SNS use in the relationship between message involvement and employees' phishing susceptibility. In collaboration with IT-Seal, an information security training company specialized in phishing simulations, we conduct a randomized field experiment with 240 organizational employees and find that phishing messages employing message involvement (i.e., pretending to be of high relevance to the recipient) yield higher phishing success, and that employees' SNS use moderates this effect. By revealing SNS users as a high-risk group for phishing attempts, our research hence provides helpful insights for information security practitioners.

Keywords: Phishing, Social networking site use, Message involvement, Randomized field experiment, Information security.

1 Introduction

Inboxes around the world have been the target of phishing attacks for decades. Phishing is an attempt to acquire private information or lure victims into opening an infected attachment through deceptive electronic communication (Wright et al., 2014a). It offers huge profit margins and ease in performing an attack, and has therefore become a prevalent problem particularly for organizations, which suffer monetary losses, substantial reputational damage or operational impacts (Wright et al., 2014a, Landesman, 2016, Piggin, 2016). Since technological threat protection measures are often ineffective, much of the responsibility to prevent phishing lies with the end user (Wright and Marett, 2010). At the same time, criminals use so-called social engineering, that is, manipulation techniques to exploit human characteristics such as habits, motives or cognitive biases (Mitnick and Simon, 2003). To counteract these incidents, it is essential to gain a deeper understanding of why exactly users fall for phish.

Phishing susceptibility falls in the research area of information security-related behavior of individuals. With 20% of research works on information security investigating phishing (Hassandoust et al., 2020), this phenomenon of deceit and deception has been explored from a variety of angles. From the attack perspective, the role of phishing email characteristics has been investigated (Williams et al., 2018, Wright et al., 2014a, Dhamija et al., 2006). Other studies have highlighted how situational factors alter the effectiveness of phishing attempts (Naidoo, 2020, Goel et al., 2017). From the user perspective, users' individual characteristics and their influence on phishing susceptibility have been explored, for example in terms of personality traits (Frauenstein and Flowerday, 2020, Moody et al., 2017), demographics (Sheng et al., 2010, Jagatic et al., 2007) or experience (Wright and Marett, 2010).

Highly targeted and context-specific attacks are referred to as spear phishing (Goel et al., 2017) and present a particularly high risk for organizational security. Spear phishers attempt to deceive employees by sending phishing emails that appear to be legitimate requests in order to elicit a certain action, such as to click on a link or divulge personal data (Schuetz et al., 2020). The attackers design email messages to be perceived of high personal relevance to the recipient, thus distracting the user in their decision making process when involved with the topic of the message (Chen et al., 2020). Experimentally, previous research has found mixed results for the effect of message involvement on phishing susceptibility amongst students (Vishwanath et al., 2011, Chen et al., 2020, Hassandoust et al., 2019, Wang et al., 2012). In this article, we therefore aim to reexamine this effect in a different context. The importance of context has been discussed extensively in IS and management research (Hong et al., 2014, Johns, 2006, Schuetz et al., 2020, Burton-Jones and Gallivan, 2007). Since spear phishers often specifically target for-profit organizations and their employees, we seek to gain a deeper understanding of the effect of message involvement by studying it in an organizational context.

Furthermore, IS research has called for further investigation of how individual differences affect users' responses to influence tactics (Pienta et al., 2020, Williams et al., 2017, Moody et al., 2017). We study phishing susceptibility through the theoretical lens of the heuristic-systematic processing model, which posits that both the fast, heuristic mode and the cognitively effortful systematic mode co-occur when being confronted with a phishing message (Goel et al., 2017, Luo et al., 2013, Vishwanath et al., 2011). In this work, we aim to respond to the above-mentioned research call by studying the phishing susceptibility of individuals particularly prone to heuristic processing, namely social networking site (SNS) users. Prior research has shown that individuals' predisposition to rely on heuristic processing is exacerbated by SNS use (Moravec et al., 2020), with SNS users being conditioned to rely on triggers and symbolic cues for online navigation and interaction (Sundar et al., 2007). Furthermore, individuals' SNS use often takes place disruptively while being engaged in other tasks, which leads to the formation of media habits in the form of mindless reactions to notifications as soon as they arrive (Vishwanath, 2016). IS research has hitherto neglected the information security behavior of SNS users with respect to phishing threats, albeit its potential for interesting interactions.

We therefore set out to investigate the following two research questions:

RQ1: How does message involvement influence employees' phishing susceptibility?

RQ2: How does the employees' social networking site (SNS) use influence this effect?

To answer our research questions, we conduct a randomized field experiment targeting 240 employees of a large German industrial manufacturing enterprise. We have chosen this methodological approach since a shortcoming of prior empirical phishing research lies in the authenticity of the captured reactions (Hassandoust et al., 2020): Previous studies often draw on surveys (Schuetz et al., 2020, Wang et al., 2017, Vishwanath et al., 2011) or "closed-lab" online experiments (Petelka et al., 2019, Wang et al., 2016, Pattinson et al., 2012, Sheng et al., 2010), while field experiments are rare. Closed-lab approaches often do not acknowledge the response to phishing threats to be a secondary task in users' daily work life, embedded in primary activities such as reading emails, and therefore are not aligned with the nature of the studied behavior (Dennis and Minas, 2018). With respect to previous field studies, those heavily rely on samples drawn from university students (Goel et al., 2017, Vishwanath, 2016, Wright et al., 2014a, Vishwanath et al., 2011, Wright and Marett, 2010, Marett and Wright, 2009). By conducting a randomized field experiment in a real-world organizational context, we address this shortcoming in our work.

Our paper offers noteworthy contributions to both research and practice. First, by studying SNS users' phishing susceptibility, we extend the research scope of how individual differences affect employees' phishing susceptibility. Second, our work offers empirical insight into actual employee susceptibility in the vulnerable context of organizational information security behavior by means of a real-world field experiment. Third, we provide valuable information for practitioners by revealing SNS users as a high-risk group for phishing attempts, hence allowing to derive concrete prevention and training measures.

2 Theoretical Background

In this section, we review pertinent literature on phishing susceptibility, addressing both the attacker and the user side. We then introduce the heuristic-systematic processing model (HSM) and compare it to other theories used in prior phishing research. Lastly, we discuss the two constructs of message involvement and SNS use to lay the theoretical foundation of this work.

2.1 Phishing susceptibility

Phishing is an attack vector used by cyber criminals (so-called phishers) in order to *"direct email recipients on a malicious link, open an infected attachment, download a piece of malware, or reply with private information"* (Wright et al., 2014a, p. 385). Cyber criminals use social engineering techniques, that is, manipulation tactics that aim at exploiting the user as the weakest link in the information security chain. These techniques often address human characteristics such as fear, curiosity or obedience to authority, or aim at exploiting habits or cognitive biases (Mitnick and Simon, 2003). Originally, phishers carried out attacks by sending rather generic messages to large numbers of recipients, and expected a small subset to be successful (so-called mass phishing). Contrary to this, spear phishing describes highly targeted, context-specific attacks that aim to appear legitimate to specific groups of individuals or organizations (Schuetz et al., 2020). Since spear phishing attacks are more sophisticated and more difficult to detect, they pose a severe risk for organizational information security. Past research suggests that after recipients click on a phishing link, they rarely detect subsequent fraud attempts such as a counterfeit login page or change their course of action (Wright and Marett, 2010). This makes the initial detection of a fraudulent email critical. Concerning the end users' point of view, Wang et al. (2016) have shown that users often exhibit a strong overconfidence in their phishing email detection abilities.

From the attack side, prior research has explored phishing emails for characteristics that make users more prone to "bite the hook". For example, Dhamija et al. (2006) have shown that visual deception tactics often succeed in fooling users, while other works have investigated the effect of influence techniques such as authority or scarcity in phishing emails (Williams et al., 2018, Wright et al., 2014a). Previous studies have devoted attention to the role of a message's content and presentation in individuals' phishing susceptibility (Wright et al., 2014b), or how contextualization increases users' likelihood of phishing victimization (Jagatic et al., 2007, Hassandoust et al., 2019, Goel et al., 2017). In the context of the COVID-19 pandemic, Naidoo (2020) has explored how situational factors (e.g., remote working or airline booking refunds) extend cybercriminals' rich assortment of deception tactics.

From the user side, numerous personal characteristics and their influence on individuals' likelihood to fall for fraudulent messages have been explored: For example, previous research works have investigated how personality traits affect individuals' susceptibility to phishing, for instance in terms of disposition to trust, curiosity or entertainment drive (Moody et al., 2017), or have explored phishing susceptibility through the lens of the Big five personality model (Frauenstein and Flowerday, 2020). Furthermore, Wright and Marett (2010) have studied how dispositional and experiential factors influence users' proneness to deception.

Demographically, prior works have found mixed evidence concerning age and gender (Li et al., 2020, Halevi et al., 2015, Sheng et al., 2010, Jagatic et al., 2007, Dhamija et al., 2006).

While extensive research has been dedicated to anti-phishing measures, for example phishing awareness training (Jensen et al., 2017, Caputo et al., 2013, Kumaraguru et al., 2008) or fear appeals (Schuetz et al., 2020), we will not elaborate on this stream of literature in this section, since it goes beyond the research scope of this work.

2.2 The Heuristic-Systematic Processing Model

Most information security behavioral studies have employed theories which assume a rational actor systematically making deliberate decisions, with 95% of them considering users' information security behavior as their primary activity on a computer (Hassandoust et al., 2020). Phishing susceptibility research in particular has used, for example, theories such as Protection Motivation Theory (Schuetz et al., 2020, Wang et al., 2017) or Technology Threat Avoidance Theory (Arachchilage and Love, 2014). However, in users' everyday work, the response to phishing threats usually happens while performing other tasks, making it cognitively difficult for users to process the situation with full attention (Dennis and Minas, 2018). The response to phishing therefore clearly is a secondary activity and should be studied as such.

A model that aligns more with the nature of phishing attacks is the Heuristic-Systematic Processing Model (HSM) (Vishwanath, 2017, Goel et al., 2017, Luo et al., 2013, Chen et al., 2020). It is closely related to the Elaboration Likelihood Model (Petty and Cacioppo, 1986) and builds on dual-process theory of information processing. Dual-process theory states that individuals process stimuli in two fundamentally different ways: (1) the central route, which involves systematic processing based on rational thinking and comparisons to prior belief; (2) the peripheral route, which focuses on simple cues, heuristics and biases (Chaiken, 1980). With regard to assessing the validity of an email, heuristic processing takes advantage of heuristic cues embedded within or surrounding the message, such as its source, length, and subject, in order to quickly make a validity assessment. In contrast, systematic processing carefully researches the message's information content with the aim to make a profound validity assessment. The HSM suggests that phishing attacks are successful either if a message can withstand systematic (System 2) processing (i.e., the victim makes a wrong assessment despite closely examining the message), or if heuristic (System 1) processing is being promoted based on false cues, which leads to quick but error-prone decisions (Luo et al., 2013). These cues can be, for example, deceptive visualization or influence tactics such as scarcity or likability, which phishers use to exploit mental shortcuts and snap judgments. Prior works on social cognition argue that heuristic processing dominates during decision-making since individuals tend to economize on mental resources (Sundar et al., 2007, Chaiken, 1980). Further research has found that the deliberate rational cognition of systematic processing is triggered only when heuristic processing detects something "that is not normal", and that when we do engage in systematic (System 2) processing, it is influenced by heuristic (System 1) processing that preceded it (Dennis and Minas, 2018). This is due to heuristic processing creating an initial impression, which subsequently seeks confirmation by systematic processing through the so-called confirmation bias. This mechanism makes System 1 cognition a powerful force in perception and behavior (Kahneman, 2011). Since heuristic processing relies on personal characteristics and experiences, heuristic processing theories are closely tied to individuals.

In this work, we build on prior works which have used the HSM and study phishing susceptibility through the theoretical lens of dual-process theory.

2.3 Message involvement

Consistent with previous research, we define message involvement as the perceived relevance of a message to the receiver (Vishwanath et al., 2011). More general, situations that individuals find involving are those that *"have intrinsic importance, personal meaning, or result in significant consequences for their lives"* (Petty and Cacioppo, 1979, p. 1916). Message involvement has hitherto predominantly been a topic of interest in advertising research. For example, the effect of message involvement on message effect or brand attitude has been explored (Wang, 2006, Laczniak et al., 1989).

In the research area of information security, prior works have begun to study how users' message involvement, when presented with a phishing email, influences their phishing susceptibility (Vishwanath et al., 2011, Chen et al., 2020, Wang et al., 2012). The findings of previous studies suggest that, on one hand, message involvement relates positively to the user's cognitive effort when processing a targeted phishing email (Wang et al., 2012), for example, with regard to the level of attention given to specific elements such as grammar and spelling (Vishwanath et al., 2011). Surprisingly, on the other hand, this does not yield a lower likelihood to fall for phish. Contrarily, Wang et al. (2012) have found no significant relationship between the user's cognitive effort in processing a phishing email and their likelihood to respond, while Vishwanath et al. (2011) have found that the level of message involvement will even increase individuals' phishing susceptibility. It has been argued that this may be due to the high level of perceived relevance distracting the user in their decision making process. While the user might process the message itself with higher cognitive effort, this might facilitate heuristic (System 1) processing with regard to the user's following action (Chen et al., 2020). Attackers can exploit this effect by using topics of personal interest as bait to lead their targets to bypass the standard phishing detection process. In our work, we reexamine the effect of message involvement on phishing susceptibility in an organizational context by raising our first research question *RQ1*.

2.4 Social networking site use

In the research field of human behavior, researchers have focused on users' behavior related to social networking sites (SNS). We refer to SNS as socio-technical systems that *"afford users the ability to engage in many forms of communication by sharing and sourcing information"* (Pike et al., 2018, p. 730). There are billions of SNS users around the world, with incumbent Facebook reaching nearly 2.5 billion monthly active users as of the end of 2019 (Kuchler et al., 2020).

Studying SNS use through the lens of dual-process theory, prior research has found that individuals' predisposition to use heuristic (System 1) processing is exacerbated by SNS use (Moravec et al., 2020). Since people generally use social media for hedonic purposes such as connecting with friends, entertainment or seeking jokes (Sledgianowski and Kulviwat, 2009), SNS users are considered to be less likely to exert cognitive System 2 effort and critically evaluate information (Moravec et al., 2020). Furthermore, they are conditioned to rely on triggers and symbolic cues for online navigation and interaction, which yields a strong reliance on heuristics (Sundar et al., 2007). It has been argued that individuals' social network behaviors may mostly be based on habit, that is, users visit SNSs regularly without conscious decision making (Chiu and Huang, 2015). The formation of a habit has been conceptualized as follows: when a person repeatedly encounters the same behavioral choice, and thus repeats their previous response to this choice, associations build up between the cues that define the context and this person's response. Given that the context remains stable, these associations then acquire a degree of automaticity (Verplanken, 2006). Individuals who routinely use SNSs unconsciously develop patterns for media use over time (Ayaburi and Andoh-Baidoo, 2019). As soon as a notification or message arrives, users with entrenched media habits tend to mindlessly react to it and click on its contents, such as a link (Vishwanath, 2016). The formation of these media habits is facilitated by the fact that SNSs are often used continually throughout the day, with triggers such as notifications presenting disruptive events while being engaged in other tasks. While the mechanisms of phishing and social engineering on social networking sites themselves have been addressed by previous works (e.g., (Vishwanath, 2017, Frauenstein and Flowerday, 2020, Algarni et al., 2017)), the role of SNS users' predispositions and media habits on phishing susceptibility has not yet been addressed in phishing research. With our second research question *RQ2*, we set out to further investigate the role of SNS use in phishing susceptibility in this work.

3 Research Model and Hypothesis Development

In this section, we develop our research model based on the Heuristic-Systematic Processing Model (Luo et al., 2013) as presented in Figure 1. We propose that message involvement has a positive effect on employees' phishing susceptibility (*H1*). Furthermore, we propose that employees' SNS use positively moderates this effect (*H2*). In the following, we develop our two hypotheses.

The message involvement that a phishing email triggers in an employee can be low or high. If the email has a topic that is impersonal, typical or expected for the employee, their message involvement will be low. In contrast, if an employee perceives the content of an email to have intrinsic importance or personal meaning, their message involvement will be high (Petty and Cacioppo, 1979). Phishers might, for example, use urgent business proposals or accuse the recipient of irregularities with their salary or vacation in order to create message involvement. Through the lens of dual-process theory, prior works have argued that, although high message involvement leads to an increase in cognitive message processing (Vishwanath et al., 2011, Wang et al., 2012), it will distract the user in their decision making process (Chen et al., 2020). This distraction will lead the user to bypass systematic processing and rely on quick, but error-prone decisions instead. The above-described relationship has been studied solely in personal or university contexts so far. In our research work, we aim to reexamine the effect in a real-world organizational context. In accordance with prior research, this leads to the following hypothesis:

H1: Message involvement has a positive effect on employees' phishing susceptibility.

In the present research context, SNS use refers to individuals' usage of social networking sites such as Facebook or LinkedIn. According to prior research works, SNS users have been found to be less likely to exert System 2 effort and critically evaluate information based on rational thinking and comparisons to prior belief (Moravec et al., 2020). Instead, they will be habituated to relying on triggers and cues, which is how most social networking sites are designed to be used (Sundar et al., 2007). Furthermore, individuals who actively use SNS will have formed entrenched media habits, and will mindlessly react to notifications or emails and its contents, such as a link (Vishwanath, 2016). In this work, we propose an interaction effect between message involvement and SNS use. In detail, we suggest that, if an employee's SNS use is low, message involvement will have a weak effect on phishing susceptibility due to the employee being distracted in their process of validating the message and relying on heuristic processing, as described above. In contrast, if an employee's SNS use is high, their media habits as well as their proneness to mindless reactions towards messages and their contents will add to their reliance on heuristic processing, and will therefore increase the likelihood of phishing susceptibility. Thus, we expect high SNS use to positively moderate the effect of message involvement on phishing susceptibility. Accordingly, we hypothesize:

H2: The employees' SNS use positively moderates the relationship between message involvement and phishing susceptibility.

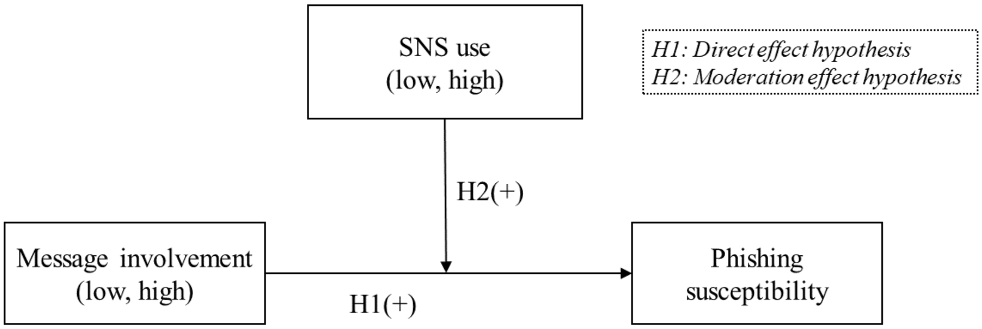

Figure 1. Research model.

4 Methodology

4.1 Research context

To date, phishing research has heavily relied on samples drawn from personal users or university students (Goel et al., 2017, Vishwanath, 2016, Wright et al., 2014a, Vishwanath et al., 2011, Wright and Marett, 2010, Marett and Wright, 2009, Wang et al., 2016, Wang et al., 2017). The reality, however, shows that for-profit organizations are particularly vulnerable to employees receiving and responding to phishing messages, since attacks on organizations (1) offer huge profit margins for cyber criminals, and (2) provide a plethora of accessible opportunities for social engineering points of attack, for example, with regard of publicly available information on the organization's management or industry. Prior information security research has found that mixed findings arise from differences between personal and organizational contexts (Schuetz et al., 2020). This can be illustrated by the simple fact that, for example, a phishing email received in a personal context presents an information security threat that targets the user's personal assets, while a phishing email received in an organizational context targets organizational information assets. Organizational threats might therefore lack personal relevance and will possibly be handled differently. Furthermore, organizational contexts have found to be rich in higher-level factors that may influence individual attitudes and behaviors (Burton-Jones and Gallivan, 2007).

We therefore argue that context matters in phishing experiments, and address this by drawing on a sample of real-world employees.

4.2 Experimental design

To verify our hypotheses as developed in Section 3 of this work, we collaborated with IT-Seal, a German information security training company. Together with IT-Seal, we conducted a randomized field experiment using a sample of 240 employees of a Germany-based enterprise of the industrial manufacturing sector. We chose this methodological approach since prior empirical information security behavior research often lacks the authenticity of the captured reactions (Hassandoust et al., 2020). We therefore aimed to align our research method with the nature of phishing: by sending simulated phishing emails to employees' inboxes during their daily work life, we were able to capture their response as a secondary activity embedded in primary activities, such as reading and responding to work-related emails.

Our collaboration partner IT-Seal offers phishing simulations for organizations to test and train their employees' security awareness. Furthermore, IT-Seal gathers data about their clients' employees' SNS use. From this data, they derive, for example, the number of SNS profiles per employee, as well as the amount of publicly shared information, such as contacts, former employees, interests or qualifications. Based on this information, IT-Seal rates the employees' SNS use as low or high. We were therefore able to classify the participants of our sample into segments of low or high SNS use. Our sample's employer was a Germany-based 2000 employees company of the industrial manufacturing sector. The participants of our study were randomly distributed across all organizational departments. Of all participants, 130 were females (54 %) and 110 were males (46 %). IT-Seal conducted the phishing simulation in close collaboration with the organization's IT department, which ensured technical deliverability of the emails by whitelisting the sender IP. Each email contained a link with a unique parameter that identified each participant.

Realistic phishing experiments rely on the use of deception, that is, the subjects are needed to not have knowledge of the existence of the study. They therefore face challenges concerning the ethical manner of the design and analysis of such an experiment (Finn and Jakobsson, 2008). In our experiment, we addressed these challenges by choosing the following four measures: (1) several weeks before the start of the phishing simulation, the organization's Chief Information Security Officer sent an email to all employees, announcing an upcoming phishing awareness training. Due to the nature of the training not being described in detail, and the announcement being sent several weeks before the phishing simulation, we posit that, at the time of receiving the simulated phishing email, our sample was not biased by their knowledge of the existence of the training. (2) Although the use of debriefing in naturalistic phishing studies is controversially discussed (Finn and Jakobsson, 2008), subjects that clicked on a link in a phishing email were debriefed by IT-Seal's landing page, which provided information on how to recognize a phishing email. Furthermore, the landing page made the employee aware of their anonymity during the training. (3) The assessment of the employees' SNS use was limited to information that was previously publicly shared by the employees themselves. (4) The phishing simulation was conducted in close accordance with the internal works council. Based on the organization's IT department's preferences, employees did not have the option to opt-out from the phishing simulation.

We conducted a 2x2 between-subject experiment with the independent variables *SNS use* (low vs. high), and *message involvement* (low vs. high) as presented in Table 1. In order to ensure a balanced sample in each group, we first segmented employees based on their SNS use. We randomly selected 120 employees from each of the two resultant segments. Afterwards, in each of these two groups, we randomly assigned employees to the control (message involvement: low) and treatment condition (message involvement: high).

		SNS use	
		Low	High
Message involvement	Low	Group A (N = 60)	Group C (N = 60)
	High	Group B (N = 60)	Group D (N = 60)

Table 1. 2x2 between-subject experimental design.

We conducted a power analysis using G*Power 3.1 (Faul et al., 2009) with the following parameter specifications: four groups (2 × 2 full-factorial design), a moderate effect size (f= 0.25), an α-level of 0.05, and a desired power level of 0.95. The results indicated that a minimum sample size of 52 per group should be sufficiently powerful to detect significant effects (Cohen, 1992, Baroudi and Orlikowski, 1989).

4.3 Manipulation of message involvement

For both the treatment and the control group, the simulated phishing email was drafted to represent a realistic spear phishing attack: both emails mimicked a legitimate email from the organization's human resources (HR) department. IT-Seal purchased domains to imitate the sender email address and internal URL (exemplarily, "...@organizatlon.com" and "https://intern.organizatlon.com/..."). For both conditions, the email contained a covered link (e.g., "click here"). The full URL could be viewed only when the employee hovered the mouse over that link. Since our sample consisted of German employees, the simulated phishing emails were sent in German language. Both simulated phishing emails were similar with regard to the length and visual impression of the email, and both email bodies consisted of plain text.

For the control condition, we chose a generic spear phishing message similar to real-world phishing attacks: the attack was tailored to the recipient as it mimicked an email coming from the company's HR department. In the email, the employee was notified that it is necessary to re-enter her "user password", followed by the instruction to click on a covered link ("You can reach the system *here*."). The subject of the email was "Set password", the sender was displayed as "Your HR team".

For the treatment condition, we also chose a phishing message mimicking a legitimate email from the company's HR department, with the alleged sender again being "Your HR team". This email, however, contained a message involvement cue. The subject of the email was "Internal job offer". The content of the email informed the recipient that she was suitable for an internal job offer, with the respective role offering more responsibility and a higher salary. The employee was instructed to click on a covered link to view the job profile ("on *this page*, you will find a PDF with the exact job description").

We chose this treatment since, according to prior research, messages that individuals find involving feature an intrinsic importance, a personal meaning, or result in significant consequences for their lives (Petty and Cacioppo, 1979). These features are given by the internal job offer phishing scenario, while at the same time retaining the context of a fraudulent internal email from the HR department. The aim of this treatment was to leverage heuristic (System 1) processing by leading the employee to rely on mental shortcuts and intuition.

Following previous experimental research (Koch and Benlian, 2017), we conducted a qualitative pre-test to assess the correct manipulation of our independent variable (i.e., message involvement). We interviewed five representatives in order to assure the realism and adequacy of the simulated phishing emails. The representatives stated that they perceived both conditions (i.e., treatment and control) as equally realistic, and found the difference between both conditions to be the perceived personal relevance, that is, message involvement. Thus, we conclude that the treatment of message involvement was manipulated as intended.

5 Analysis and Results

Before analyzing the results, we conducted several one-way ANOVAs to assess whether we successfully randomly assigned employees to the four groups. The results indicate that we succeeded in our distribution as employees' demographics and controls did not significantly confound the effects of our manipulation.

In total, IT-Seal sent one simulated phishing email to each of the 240 employees. Records showed that 63 employees clicked on the link embedded in the phishing message, resulting in a total phishing susceptibility of 26.3 percent. Table 1 summarizes the descriptive statistics of employees' phishing susceptibility in each of the four groups.

Group	Click rate	Male	Female
A Low message involvement and low SNS use	13.33%	24	36
B High message involvement and low SNS use	25.00%	30	30
C Low message involvement and high SNS use	15.00%	30	30
D High message involvement and high SNS use	51.66%	26	34

Table 2. Descriptive statistics by group.

To test our hypotheses, we conducted a hierarchical regression analysis with SPSS 27. Table 2 shows the three hierarchical regressions we estimated for the dependent variable (i.e., phishing susceptibility): including only the control variable (Model 1), adding message involvement and SNS use (Model 2), and adding the interaction of message involvement with SNS use in the third step (Model 3). The results of the first model (Model 1) revealed that the control variable was not significant ($p > 0.05$). The results of the second model (Model 2) indicated that the effect of message involvement on phishing susceptibility was positive and significant ($\beta = 1.33$, $p < 0.001$). This demonstrates that message involvement, that is, perceived personal relevance in a phishing email, increases the likelihood to be phished, **supporting H1**. The results of the last model (Model 3) revealed that the interaction of message involvement and SNS use had a positive significant effect (β = 1.31, $p < 0.05$) on phishing susceptibility. Thus, SNS use moderates the relationship between message involvement and phishing susceptibility, **supporting H2**.

	Model 1		**Model 2**		**Model 3**	
	Coefficient	S.E.	Coefficient	S.E.	Coefficient	S.E.
Control						
Gender	-0.36	0.33	-0.25	0.32	-0.15	0.31
Dependent variables						
Message involvement			1.33***	0.33	0.63	0.46
SNS use			0.78**	0.32	-0.13	0.52
Message involvement × SNS use					1.31*	0.36
Model Fit						
Log Likelihood	-274.85		-252.39		-248.44	
Nagelkerke R^2	0.01		0.16		0.18	
Omnibus model χ^2	1.49		23.92***		27.87***	
Notes: *p < 0.05; **p < 0.01; ***p < 0.001, N = 240						

Table 3. Binary logistic regression analysis on phishing susceptibility.

To facilitate the interpretation of the findings, we followed the procedures of (Aiken et al., 1991), and plotted the moderation effect of SNS use in Figure 2. We consistently observed an interaction effect, such that employees with high SNS use are more susceptible to the treatment. However, a significant difference in phishing susceptibility between employees with low SNS use and employees with high SNS use did not emerge when the phishing email did not include message involvement cues.

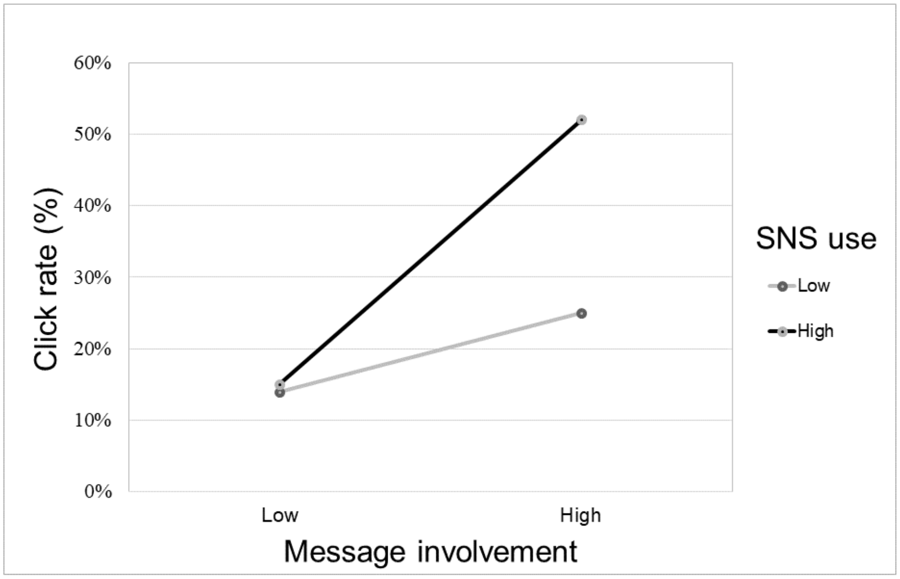

Figure 2. Moderating effect of SNS use on the relationship between message involvement and phishing susceptibility.

6 Discussion

With our study, we aimed to investigate the effects of message involvement and SNS use on employees' phishing susceptibility in a real-world organizational context. Building on the Heuristic Systematic Model, we proposed a conceptual model in which message involvement exerts influence on phishing susceptibility. Furthermore, our research considered the moderating effect of SNS use on the relationship between message involvement and phishing susceptibility. Our empirical findings are not only consistent with those of prior studies on message involvement (Chen et al., 2020, Vishwanath et al., 2011), but also offer new insights. We find that SNS use positively moderates the effect of message involvement on phishing susceptibility: As proposed, when SNS use is high, message involvement has a strong effect on phishing susceptibility. We argue that this is due to SNS users' proneness to heuristic processing interacting with the distraction caused by message involvement. Our study thus provides not only theoretical contributions but also crucial practical implications for organizations' security practitioners.

6.1 Theoretical contributions

Our work contributes to research in several ways. First, our study answers IS research calls for further investigation of how individual differences affect users' responses to influence tactics (Moody et al., 2017, Williams et al., 2017, Pienta et al., 2020). We extend prior research on the role of users' characteristics in their susceptibility to phishing attacks by studying the influence of their SNS use. Drawing on the Heuristic Systematic Model, we have shown that individuals' SNS use, which comes along with a predisposition to rely on heuristic (System 1) processing, positively moderates the effect of high message involvement on phishing susceptibility. For low message involvement, however, SNS use does not affect users' likelihood to fall for phish. This reveals that users' information processing in the context of phishing is highly complex, and that future research is needed to better understand the underlying mechanisms.

Second, while previous research has studied the phenomenon of phishing from both the attack and the user side, a crucial shortcoming of prior studies lies in the research context. Whereas the importance of context has been discussed in information security research (Schuetz et al., 2020), prior IS research works in the field of phishing susceptibility heavily rely on samples drawn from personal users or university students (Goel et al., 2017, Vishwanath, 2016, Wright et al., 2014a, Vishwanath et al., 2011, Wright and Marett, 2010, Marett and Wright, 2009, Wang et al., 2016, Wang et al., 2017). It is, however, for-profit organizations that are particularly vulnerable to phishing attacks and their consequences. By using a sample drawn from a German industrial manufacturing company, we were able to study real-world employees' behavior. Our work thus extends the research scope of increasing literature on phishing by the important and unique context of real-world organizations.

Thirdly, our research addresses the shortcoming of empirical insight into actual user behavior in IS research. Prior studies commonly lack the authenticity of the captured reactions, since the experimental setup does not align with the nature of phishing (Hassandoust et al., 2020). In our work, we have conducted a randomized field experiment by sending simulated phishing emails to real-world employees during their everyday work, as well as directly measuring their SNS use based on the information they publicly shared online. This approach offers insights into actual employee security behavior, and highlights the role of collaboration partners, such as security training providers, in IS research.

6.2 Implications for practice

Our research also outlines important practical implications. We have shown that message involvement generates a higher phishing vulnerability in the organizational context. Prior research has found that influence techniques lose some of their potency when the targets are aware of the techniques and their susceptibility to them (Wright et al., 2014a). Our work therefore helps to raise awareness of phishing influence techniques among users, thereby reducing the effectiveness of phishing attacks both in the personal and organizational context.

Furthermore, our results provide valuable insight for information security officers and phishing training providers. Understanding the moderating role of SNS use in employees' phishing vulnerability allows to derive concrete training and prevention measures. Such concrete measures might replace the one-size-fits-all approach taken by most security trainings to date, yielding for training measures that are more effective to prevent phishing incidents. Moreover, our study has highlighted that SNS users are particularly prone to be deceived by phishing emails using certain influence tactics. This reveals a combustible mix, since SNS users are commonly those individuals who publicly share extensive information about themselves and their organizations online, which in turn provides a breeding ground for social engineering attacks: Phishers can easily use this information to tailor targeted spear phishing messages in the knowledge that their victims will easily get hooked. Therefore, SNS users should be considered a high-risk group regarding phishing attacks, and implementing measures to increase their security awareness as well as their phishing detection capabilities is crucial.

6.3 Limitations and future research

Despite the above contributions, some limitations of our study provide fruitful opportunities for future research. First, our study is conducted in the context of a German organization. Since users' security behavior might be subject to cultural differences (Butavicius et al., 2017), further research is needed to examine the robustness of our findings in culturally diverse environments (e.g., by conducting studies with culturally different samples). Furthermore, our sample consisted of employees of the industrial manufacturing sector, leaving room for the exploration of industrial differences between organizations. It can be argued that an organization's industry will influence employees' security behavior, particularly with regard to the role of information technology within the industry. We hope that further studies will be conducted to compare the behaviors of employees from different industries.

Second, while our sample consisted of real employees that are representative of the general organizational user population, it also limited our control over the experimental procedure. During the experiment, our only measuring points were if a simulated phishing email was sent, received, and if the contained link was clicked. It is therefore possible that the participation of some subjects was not entirely according to the experimental design. Furthermore, the design of or study did not allow for capturing and controlling for many individual factors germane to understanding individuals' susceptibility to phishing attacks. We suggest future research to employ complementary research methods (e.g., surveys or lab experiments) in order to improve our understanding of SNS users' phishing susceptibility.

Third, while we measured our participants' SNS use directly (in contrast to, e.g., self-reported behavior), we were not able to monitor the employees' involvement with or activity on SNS at the exact same time as the phishing simulation was conducted. On the basis of the HSM, we believe that investigating the effect of, for example, parallel use of SNS and email on phishing susceptibility provides an interesting opportunity for future research.

7 Conclusion

Phishing is a threat not only to organizations' security, but also to their profitability and reputation. By performing this study, we cast light on why employees fall for phishing attempts. We have extended the research scope of previous phishing literature by studying phishing vulnerability in a real-world organizational context. Furthermore, we have synthesized research on social networking site use and phishing, and built a research model that incorporates SNS use in phishing vulnerability by drawing on dual-process theory. By means of a randomized field experiment, we have provided evidence for the moderating role of SNS use on the relationship between message involvement and phishing susceptibility.

Insights from this study can help organizations understand their vulnerability towards phishing attacks, and guide individualized prevention approaches.

Acknowledgement

We thank our collaboration partner IT-Seal for the opportunity to conduct this field experiment together. We would also like to express our gratitude to the anonymous reviewers of this paper, whose guidance and suggestions were very valuable to improve the lucidity and quality of this work.

References

Aiken, L. S., West, S. G. & Reno, R. R. (1991). *Multiple regression: Testing and interpreting interactions.* SAGE Publications.

Algarni, A., Xu, Y. & Chan, T. (2017). "An empirical study on the susceptibility to social engineering in social networking sites: the case of Facebook." *European Journal of Information Systems*, 26 (6), pp. 661-687.

Arachchilage, N. a. G. & Love, S. (2014). "Security awareness of computer users: A phishing threat avoidance perspective." *Computers in Human Behavior*, 38, pp. 304-312.

Ayaburi, E. & Andoh-Baidoo, F. K. (2019). "Understanding phishing susceptibility: an integrated model of cue-utilization and habits."

Baroudi, J. J. & Orlikowski, W. J. (1989). "The problem of statistical power in MIS research." *MIS quarterly*, 13 (1), pp. 87-106.

Burton-Jones, A. & Gallivan, M. J. (2007). "Toward a deeper understanding of system usage in organizations: a multilevel perspective." *MIS quarterly*, 31 (4), pp. 657-679.

Butavicius, M. A., Parsons, K., Pattinson, M. R., Mccormac, A., Calic, D. & Lillie, M. (2017). "Understanding susceptibility to phishing emails: Assessing the impact of individual differences and culture." *Human Aspects of Information Security & Assurance*, Adelaide, Australia.

Caputo, D. D., Pfleeger, S. L., Freeman, J. D. & Johnson, M. E. (2013). "Going spear phishing: Exploring embedded training and awareness." *IEEE Security & Privacy*, San Francisco, USA. pp. 28-38.

Chaiken, S. (1980). "Heuristic versus systematic information processing and the use of source versus message cues in persuasion." *Journal of personality and social psychology*, 39 (5), pp. 752.

Chen, R., Gaia, J. & Rao, H. R. (2020). "An examination of the effect of recent phishing encounters on phishing susceptibility." *Decision Support Systems*, 133.

Chiu, C.-M. & Huang, H.-Y. (2015). "Examining the antecedents of user gratification and its effects on individuals' social network services usage: the moderating role of habit." *European Journal of Information Systems*, 24 (4), pp. 411-430.

Cohen, J. (1992). "A power primer." *Psychological bulletin*, 112 (1), pp. 155.

Dennis, A. R. & Minas, R. K. (2018). "Security on autopilot: Why current security theories hijack our thinking and lead us astray." *ACM SIGMIS Database: the DATABASE for Advances in Information Systems*, 49 (SI), pp. 15-38.

Dhamija, R., Tygar, J. D. & Hearst, M. (2006). "Why phishing works." *CHI Conference on Human Factors in Computing Systems - CHI '06*

Montréal, Canada. Montréal, Québec, Canada, pp. 581-590.

Faul, F., Erdfelder, E., Buchner, A. & Lang, A.-G. (2009). "Statistical power analyses using G* Power 3.1: Tests for correlation and regression analyses." *Behavior research methods*, 41 (4), pp. 1149-1160.

Finn, P. & Jakobsson, M. (2008). "Designing and Conducting Phishing Experiments." *IEEE Technology and Society Magazine, Special Issue on Usability and Security*, 26 (1), pp. 46-58.

Frauenstein, E. D. & Flowerday, S. (2020). "Susceptibility to phishing on social network sites: A personality information processing model." *Computers & Security*, 94.

Goel, Goel, S., Williams, K., University at Albany, S., Dincelli, E. & University at Albany, S. (2017). "Got Phished? Internet Security and Human Vulnerability." *Journal of the Association for Information Systems*, 18 (1), pp. 22-44.

Halevi, T., Memon, N. & Nov, O. (2015). "Spear-Phishing in the Wild: A Real-World Study of Personality, Phishing Self-Efficacy and Vulnerability to Spear-Phishing Attacks." *SSRN Electronic Journal*.

Hassandoust, F., Singh, H. & Williams, J. (2019). "How Contextualisation Affects the Vulnerability of Individuals to Phishing Attempts." *Pacific Asia Conference on Information Systems*, Xi'an, China.

Hassandoust, F., Techatassanasoontorn, A. A. & Singh, H. (2020). "Information Security Behaviour: A Critical Review and Research Directions." *European Conference on Information Systems*, Virtual conference.

Hong, W., Chan, F. K., Thong, J. Y., Chasalow, L. C. & Dhillon, G. (2014). "A framework and guidelines for context-specific theorizing in information systems research." *Information Systems Research*, 25 (1), pp. 111-136.

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

Jensen, M. L., Dinger, M., Wright, R. T. & Thatcher, J. B. (2017). "Training to Mitigate Phishing Attacks Using Mindfulness Techniques." *Journal of Management Information Systems*, 34 (2), pp. 597-626.

Johns, G. (2006). "The essential impact of context on organizational behavior." *Academy of management review*, 31 (2), pp. 386-408.

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

Koch, O. F. & Benlian, A. (2017). "The effect of free sampling strategies on freemium conversion rates." *Electronic Markets*, 27 (1), pp. 67-76.

Kuchler, T., Russel, D. & Stroebel, J. (2020). "The geographic spread of COVID-19 correlates with structure of social networks as measured by Facebook." National Bureau of Economic Research.

Kumaraguru, P., Sheng, S., Acquisti, A., Cranor, L. F. & Hong, J. (2008). "Lessons from a real world evaluation of anti-phishing training." *2008 eCrime Researchers Summit*, Atlanta, USA.

Laczniak, R. N., Muehling, D. D. & Grossbart, S. (1989). "Manipulating message involvement in advertising research." *Journal of advertising*, 18 (2), pp. 28-38.

Landesman, T. (2016). *55 Companies and Counting - W-2 Spear Phishing Attacks Continue to Increase* [Online]. Available: https://blog.cloudmark.com/2016/03/31/55-companies-and-counting-w-2-spear-phishing-attacks-continue-to-increase/ [Accessed 08.07.2020].

Li, W., Lee, J., Purl, J., Greitzer, F. L., Bahram, Y. & Laskey, K. (2020). "Experimental Investigation of Demographic Factors Related to Phishing Susceptibility." *Hawaii International Conference on System Sciences*, Maui, USA.

Luo, X., Zhang, W., Burd, S. & Seazzu, A. (2013). "Investigating phishing victimization with the Heuristic–Systematic Model: A theoretical framework and an exploration." *Computers & Security*, 38, pp. 28-38.

Marett, K. & Wright, R. (2009). "The Effectiveness of Deceptive Tactics in Phishing." *Americas Conference on Information Systems*, San Francisco, USA. pp. 10.

Mitnick, K. D. & Simon, W. L. (2003). *The art of deception: Controlling the human element of security.* John Wiley & Sons.

Moody, G. D., Galletta, D. F. & Dunn, B. K. (2017). "Which phish get caught? An exploratory study of individuals′ susceptibility to phishing." *European Journal of Information Systems*, 26 (6), pp. 564-584.

Moravec, P. L., Kim, A. & Dennis, A. R. (2020). "Appealing to Sense and Sensibility: System 1 and System 2 Interventions for Fake News on Social Media." *Information Systems Research*, 31 (3), pp. 987-1006.

Naidoo, R. (2020). "A multi-level influence model to COVID-19 themed cybercrime." *European Journal of Information Systems*, 29 (3), pp. 306-321.

Pattinson, M., Jerram, C., Parsons, K., Mccormac, A. & Butavicius, M. (2012). "Why do some people manage phishing e-mails better than others?" *Information Management & Computer Security*, 20 (1), pp. 18-28.

Petelka, J., Zou, Y. & Schaub, F. (2019). "Put Your Warning Where Your Link Is: Improving and Evaluating Email Phishing Warnings." *CHI Conference on Human Factors in Computing Systems - CH '19*, Glasgow, UK.

Petty, R. E. & Cacioppo, J. T. (1979). "Issue involvement can increase or decrease persuasion by enhancing message-relevant cognitive responses." *Journal of personality and social psychology*, 37 (10), pp. 1915.

Petty, R. E. & Cacioppo, J. T. (1986). *Communication and Persuasion: Central and Peripheral Routes to Attitude Change.* New York, NY: Springer.

Pienta, D., Thatcher, J. B. & Johnston, A. (2020). "Protecting a whale in a sea of phish." *Journal of Information Technology*, 35 (3), pp. 214-231.

Piggin, R. (2016). "Cyber security trends: What should keep CEOs awake at night." *International Journal of Critical Infrastructure Protection*.

Pike, J. C., Bateman, P. J. & Butler, B. S. (2018). "Information from social networking sites: Context collapse and ambiguity in the hiring process." *Information Systems Journal*, 28 (4), pp. 729-758.

Schuetz, S. W., Lowry, P. B., Pienta, D. A. & Thatcher, J. B. (2020). "The Effectiveness of Abstract Versus Concrete Fear Appeals in Information Security." *Journal of Management Information Systems (accepted 14-May-2020).*

Sheng, S., Holbrook, M., Kumaraguru, P., Cranor, L. F. & Julie, D. (2010). "Who falls for phish? a demographic analysis of phishing susceptibility and effectiveness of interventions." *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, Atlanta, Georgia, USA. Association for Computing Machinery, pp. 373–382.

Sledgianowski, D. & Kulviwat, S. (2009). "Using social network sites: The effects of playfulness, critical mass and trust in a hedonic context." *Journal of computer information systems*, 49 (4), pp. 74-83.

Sundar, S. S., Knobloch-Westerwick, S. & Hastall, M. R. (2007). "News cues: Information scent and cognitive heuristics." *Journal of the American society for information science and technology*, 58 (3), pp. 366-378.

Verplanken, B. (2006). "Beyond frequency: Habit as mental construct." *British Journal of Social Psychology*, 45 (3), pp. 639-656.

Vishwanath, A. (2016). "Mobile device affordance: Explicating how smartphones influence the outcome of phishing attacks." *Computers in Human Behavior*, 63, pp. 198-207.

Vishwanath, A. (2017). "Getting phished on social media." *Decision Support Systems*, 103, pp. 70-81.

Vishwanath, A., Herath, T., Chen, R., Wang, J. & Rao, H. R. (2011). "Why do people get phished? Testing individual differences in phishing vulnerability within an integrated, information processing model." *Decision Support Systems*, 51 (3), pp. 576-586.

Wang, Wang, J., Li, Y., Columbia, C., Rao, H. R. & The University of Texas at San, A. (2016). "Overconfidence in Phishing Email Detection." *Journal of the Association for Information Systems*, 17 (11), pp. 759-783.

Wang, A. (2006). "Advertising engagement: A driver of message involvement on message effects." *Journal of Advertising Research*, 46 (4), pp. 355-368.

Wang, J., Herath, T., Chen, R., Vishwanath, A. & Rao, H. R. (2012). "Research article phishing susceptibility: An investigation into the processing of a targeted spear phishing email." *IEEE transactions on professional communication*, 55 (4), pp. 345-362.

Wang, J., Li, Y. & Rao, H. R. (2017). "Coping Responses in Phishing Detection: An Investigation of Antecedents and Consequences." *Information Systems Research*, 28 (2), pp. 378-396.

Williams, E. J., Beardmore, A. & Joinson, A. N. (2017). "Individual differences in susceptibility to online influence: A theoretical review." *Computers in Human Behavior*, 72, pp. 412-421.

Williams, E. J., Hinds, J. & Joinson, A. N. (2018). "Exploring susceptibility to phishing in the workplace." *International Journal of Human-Computer Studies*, 120, pp. 1-13.

Wright, R. T., Jensen, M. L., Thatcher, J. B., Dinger, M. & Marett, K. (2014a). "Influence Techniques in Phishing Attacks: An Examination of Vulnerability and Resistance." *Information Systems Research*, 25 (2), pp. 385-400.

Wright, R. T. & Marett, K. (2010). "The Influence of Experiential and Dispositional Factors in Phishing: An Empirical Investigation of the Deceived." *Journal of Management Information Systems*, 27 (1), pp. 273-303.

Wright, R. T., Marett, K. & Thatcher, J. B. (2014b). "Extending Ecommerce Deception to Phishing." *International Conference on Information Systems*, Auckland, New Zealand.