Two Essays on Antiphishing Performance: Impacts of Task Performance, Feedback Intervention and Phishing Characteristics

By

Shihe Pan

A dissertation submitted in partial fulfillment of

the requirements for the degree of

Doctor of Philosophy

(Business)

at the

UNIVERSITY OF WISCONSIN-MADISON

2022

Date of final oral examination: 7/28/2022

The dissertation is approved by the following members of the Final Oral Committee:
Sung S. Kim, Professor, Business (Advisor)
Terence Ow, Professor, Business
Qinglai He, Assistant Professor, Business
Jungwon Kuem, Assistant Professor, Business
Alan Rubel, Professor, Information

© Copyright by Shihe Pan 2022

All Rights Reserved

Acknowledgments

I sincerely thank my advisor, Sung Kim, for his continuous support during my Ph.D. study. The discussions between us inspired me to think about my research as well as life. I am also very thankful to Dr. Jungwon Kuem and Dr. Austin Kwak, for their invaluable advice and guidance throughout this research project. I also highly appreciate the help from the Office of Cybersecurity at the University of Wisconsin-Madison with the data collection. Further, I would like to thank all the colleagues in the Department of Operations and Information Management for all their comments on my research. Finally, I cannot forget to thank my family and friends for their love and support.

Table of Contents

Dissertation Abstract vi

Essay 1: The Impact of Attention and Task Performance on Antiphishing Performance: A Multigoal Perspective 1

- Abstract 1
- 1. Introduction 2
- 2. Theoretical Framework 5
  - 2.1. Prior Research on Antiphishing 5
  - 2.2. Model of Safety Performance 7
  - 2.3. Antiphishing as a Goal in Daily Work 9
- 3. Research Model and Hypotheses 11
- 4. Research Method 19
  - 4.1. Sample and Data Collection 19
  - 4.2. Measures 21
- 5. Analyses and Results 24
  - 5.1. Data Analyses 24
  - 5.2. Preliminary Results 24
  - 5.3. Exploratory Factor Analysis 25
  - 5.4. Confirmatory Factor Analysis 25
  - 5.5. Structural Equation Models 28
  - 5.6. Generalized Estimating Equations 32
- 6. Discussion and Conclusion 37
  - 6.1. Theoretical Contributions 38
  - 6.2. Practical Implications 40
  - 6.3. Limitations and Future Research 41
  - 6.4. Conclusions 42
- References 43
- Appendix A: Prior Studies on Antiphishing Performance 49
- Appendix B: Legitimate and Phishing Emails 52

Essay 2: Roles of Feedback and Phishing Characteristics in Identifying Phishing Scams: Conceptual Model and Three Experiments 55

- Abstract 55
- 1. Introduction 56
- 2. Theoretical Background 59
  - 2.1. Feedback and Phishing Characteristics 59
  - 2.2. Theories of Feasibility 61
  - 2.3. Skill Acquisition Theory 62
- 3. Research Model and Hypotheses 64
  - 3.1. Feedback Type and Perceived Detection Feasibility 64
  - 3.2. Feedback Quantity and Phishing Cue Saliency 67
- 4. Methods 69
  - 4.1. Experiment 1 71
  - 4.2. Experiment 2 74

  4.3. Experiment 3 79
5. Discussion and Conclusion 84
  5.1. Summary of Findings 84
  5.2. Theoretical Contributions 86
  5.3. Practical Contributions 88
  5.4. Limitations and Future Research Directions 89
  5.5. Concluding Remarks 90
References 90
Appendix C: Measurement Items 96
Appendix D: Email Samples and Manipulations 98
Appendix E: Descriptive Statistics, Composite Reliabilities, AVEs, and Correlation Matrix 100

A List of Tables

Table 1. Measurement Items………………………………………………………………………… 22

Table 2. Clicking Rates by Gender……………………………………………………………….. 25

Table 3. Item Loadings for Key Research Constructs……………………………………………….. 26

Table 4. Means, Standard Deviations and Correlations of Measurement Scales………………………….. 27

Table 5. Statistics for Internal Consistency and Reliability……………………………………………….. 28

Table 6. Results of Structural Equation Modeling………………………………………………….... 30

Table 7. Results of Generalized Estimating Equations…………………………………………….... 34

Table 8. Prior Studies on Antiphishing Performance…………………………………………………... 49

Table 9. Summary of Experiments……………………………………………………………….... 70

Table 10. Experiment 1: ANCOVA Results for Decision Avoidance and Detection Accuracy………….. 73

Table 11. Experiment 2: GEE Results for Decision Avoidance and Detection Accuracy………………… 77

Table 12. Experiment 3: GEE Results for Decision Avoidance and Detection Accuracy………………… 82

Table 13. Summary of Research Findings………………………………………………………… 85

Table 14. Experiment 2……………………………………………………………………………….. 100

Table 15. Experiment 3……………………………………………………………………………... 101

A List of Figures

Figure 1. Model of Safety Performance………………………………………………………………. 8
Figure 2. Goal-Oriented Framework…………………………………………………………………….. 9
Figure 3. Conflicting Goals……………………………………………………………………………….. 11
Figure 4. Conceptual Model………………………………………………………………………………. 12
Figure 5. Results of Structural Equation Modeling………………………………………………….. 31
Figure 6. 3-D Plots of GEE Results…………………………………………………….......................... 35
Figure 7. Plots of the Interaction Effect……………………………………………............................. 36
Figure 8. Research Model………………………………………………... 64
Figure 9. Experimental Procedure………………………………………………………………………... 71
Figure 10. Experiment 2: Interaction between Feedback Type and Perceived Detection Feasibility……… 78
Figure 11. Experiment 3: Interaction between Feedback Quantity and Phishing Cue Saliency…………… 84

Dissertation Abstract

Phishing is a prevalent cyberattack which can cause huge financial losses. With the exception of existing automated tools to defend against phishing attacks, individual efforts are also recognized as an important part in combating these attacks. This dissertation consists of two essays investigating the key factors influencing individuals' antiphishing performance (i.e., performance on detecting phishing emails) and exploring possibilities to assist individuals in improving their antiphishing performance. Essay 1 aims to develop and test a new model of antiphishing performance in a multigoal context. The study tested the model by using data collected from 357 participants through a field experiment in which four legitimate emails and four phishing emails were sent over a period of five weeks and a post-survey at the end of the experiment. Essay 2 examines the relative effectiveness of different feedback materials for improving individuals' antiphishing performance in antiphishing training. It also examines the interaction between feedback and phishing characteristics that influence antiphishing performance. To understand the impacts of feedback and phishing characteristics, we conducted three experiments with 514 subjects in the United States.

The dissertation provides several insights into individuals' antiphishing performance. First, individuals' task performance in daily work (e.g., keeping up with work emails) is negatively associated with antiphishing performance for task-related phishing emails. Attention to phishing cues is positively associated with antiphishing performance, and it also weakens the negative relationship between task performance and antiphishing performance. Second, feedback with technical information is more effective than feedback with abstract information in improving individuals' antiphishing performance during antiphishing training. Moreover, the quantity of the technical feedback interacts with phishing characteristics (i.e, phishing cue saliency) to influence antiphishing performance.

In summary, the dissertation has theoretical contributions as well as practical implications. The findings facilitate our understanding regarding individuals' antiphishing performance and provide helpful guidelines on the design of antiphishing training programs.

Essay 1: The Impact of Attention and Task Performance on Antiphishing Performance: A Multigoal Perspective

Abstract

People are often recognized as an important part in thwarting phishing attacks. However, the goal of thwarting phishing can conflict with other goals at work, and we still lack an understanding of this multigoal scenario. The objective of this study is to develop and test a new model of antiphishing performance, especially in relation to individuals' daily task performance. Drawing on the model of safety performance and a goal-oriented framework, we developed a model that highlights the important roles of antiphishing self-efficacy, antiphishing climate, antiphishing motivation, attention to phishing cues, and task performance in explaining antiphishing behavior. We tested the model by using a field experiment with a survey questionnaire in which four legitimate work emails and four phishing emails were sent to 357 participants. The data were analyzed through structural equation modeling and generalized estimating equations. As hypothesized, antiphishing climate, antiphishing self-efficacy, and antiphishing motivation played important roles in regulating attention to phishing cues. The results also showed that task performance in keeping up with work emails was negatively related with antiphishing performance for task-related phishing emails. Attention to phishing cues was positively associated with antiphishing performance; it also decreased the negative relationship between task performance and antiphishing performance for task-related phishing emails. These findings shed light on the potential of a goal-oriented framework in studying antiphishing behavior and have practical implications that bear on the conflict between antiphishing performance and task performance.

Keywords: phishing, goal-oriented, task performance, antiphishing performance, attention to phishing cues

1. Introduction

Phishing is an effort to trick individuals into providing sensitive information via electronic communication, usually through emails mimicking a trusted source. Successful attacks can result in not only loss of individual data but also have detrimental consequences for businesses (Hong, 2012). Such successful attacks have an average annual cost of 1.4 million dollars per organization (Bissell et al., 2019). Because phishing emails often exploit a vulnerability in human psychology (Goel et al., 2017; Wright, 2014), people are often recognized as an important part in the process of thwarting these attacks (Arachchilage & Love, 2014; Jensen et al., 2017; Sheng et al., 2007). Nevertheless, people often pursue multiple goals in their daily work (e.g., performing multiple work- and security-related activities). Specifically, for example, a market analyst may have a goal of completing a sales report by a deadline while having at the same time another goal of fulfilling the learning requirements of the company's regular security education training program. A student may focus on a goal of quickly finishing her assignments while having another goal of checking the incoming emails attentively to avoid phishing messages. These goals can conflict with each other if the pursuit of one detracts or inhibits the pursuit of another because of limited physical or attentional resources (Vancouver et al., 2010). Such conflicts may exist especially between goals of fending off phishing attacks and completing other work-related tasks. Similar to the prior student example, individuals may check their daily emails hastily and succumb to phishing attacks accidentally as they also hurry to achieve their primary goals by the required deadlines. This situation calls for studies about how people perform in thwarting phishing attacks while trying to achieve other potentially conflicting goals.[1]

Prior research has yielded many insights into people's antiphishing behavior (i.e., dealing with phishing emails) and performance. For example, one stream of research has examined different factors influencing antiphishing behavior and performance. Among those factors are self-efficacy, risk perception, negative emotions, coping responses, habitual usage, and prior phishing experience (Vishwanath, 2015; Wang et al., 2017; Wright & Marett, 2010). Another stream of research has focused on developing effective antiphishing training programs or warning systems to promote appropriate antiphishing behavior and high antiphishing performance. Studies in this stream have evaluated the effectiveness of embedded training programs (Kumaraguru et al., 2007a; Kumaraguru et al., 2007b), interactive educational games (Arachchilage & Love, 2014; Dincelli & Smith, 2020; Sheng et al., 2007), mindfulness (Jensen et al., 2017) and also different warning systems (Abbasi et al., 2021; Nguyen et al., 2021). The third stream of research has focused on the effectiveness of phishers' strategies for deceiving people. Studies in this stream have compared the effectiveness of different deceptive email strategies such as using fear of loss and authority (Goel et al., 2017; Wright, 2014).

---

[1] We used the word "work" in this paper rather generally as a certain task (e.g., study, shopping, performing organizational duties) but not exclusively as activities for earning a living as an employee in an organization. Accordingly, we use "task" and "work" interchangeably in the paper.

The current study aims to address a research gap in the first stream of phishing research (i.e., antecedents of antiphishing behavior and performance). Specifically, thwarting phishing emails is one of many goals in individuals' lives. Few studies have explored how individuals' task performance in their daily work influences their antiphishing performance despite the existence of goal conflicts between daily work-related goals and the goal of combating phishing emails. Although little research on antiphishing behavior has paid attention to a multigoal situation, a number of prior studies examining security policy compliance and human cybersecurity behavior have recognized a potential conflict between work- and security-related goals (Bulgurcu et al., 2010; Chowdhury et al., 2018; Mayer et al., 2017). Moreover, the goal of combating phishing emails is probably a subordinate goal for most people, which is superseded by other higher order, focal goals (e.g., accomplishing high work productivity in general). Thus, we believe it is necessary to understand how people respond to phishing emails while dealing with other potentially conflicting goals.

In addition, prior research discussed the influence of individual characteristics such as prior experiences, self-efficacy, motivation, and emotion on antiphishing behavior and performance but ignored the influence of the antiphishing climate. Individuals' goal-driven behavior and goal attainment in combating phishing emails are influenced by the antiphishing climate as well as by their individual characteristics. The goal of combating phishing emails also conflicts with their pursuits of other work-related goals. Thus, we believe combining the goal-oriented framework and the antiphishing climate to explain individuals' goal-driven behavior and goal attainment in combating phishing emails will provide a more comprehensive view of their antiphishing behavior and performance.

The objective of this study is to develop and test a new model of antiphishing performance (i.e., individuals' performance in thwarting phishing emails), especially in relation to individuals' daily task performance. Drawing on the model of safety performance and a goal-oriented framework, we developed a model that explains antiphishing behavior when people also have a goal of keeping up with work emails. We especially extended the model of safety performance in a traditional offline work environment to explain antiphishing performance. Specifically, we suggested antecedents of antiphishing performance that include not only well-known factors such as antiphishing self-efficacy but also less-studied factors such as antiphishing climate, motivation, and attention to phishing cues. Rooted in goal theories, our framework also suggests a negative relationship between antiphishing performance and task performance in keeping up with work emails as well as the moderating role of attention in mitigating this negative relationship. The proposed model was evaluated through a field experiment with a survey including objective behavioral responses of 357 participants to eight legitimate and phishing emails over five weeks in a real-life setting.

Our study contributes significantly to information systems (IS) research in several ways. First, we are among the first to show the promise of the goal-oriented framework as a conceptual tool for understanding antiphishing behavior. Similarly, this study is the first to demonstrate the correspondence of this framework to a specific model of safety performance that suggests antecedents of antiphishing performance such as climate, self-efficacy, motivation, and attention. Second, we have predicted and confirmed a counterintuitive result that shows that an antiphishing climate is negatively related to attention to phishing cues. This finding reveals a boundary condition of the safety model that was originally designed for a traditional offline work environment and extends it by highlighting the dynamically evolving nature of phishing attempts. Third, we have demonstrated that a negative relationship prevails between antiphishing performance and task performance in keeping up with work emails. The present finding highlights the importance of evaluating antiphishing not only in its own right but also within a larger context encompassing other conflicting goals. Finally, the proposed model suggests a moderating role of attention between task performance and antiphishing performance. The finding has potential to shed a valuable light on an effective antiphishing approach that can potentially reduce the negative impact of task performance on antiphishing performance.

In the remaining part of this paper, we first have illustrated the theoretical background of our study. We then presented our research model and elaborated on our hypotheses. Following the hypotheses development are the Method section and the Analyses and Results section. At the end of the paper, we have discussed the contributions and Limitations of our study.

2. Theoretical Framework

2.1. Prior Research on Antiphishing

A stream of prior research has examined experiential, dispositional, emotional, and demographic factors that influence antiphishing behavior and performance. For example, Wright and Marett (2010) studied the impact on antiphishing behavior of computer self-efficacy, security knowledge, web experience, trust, perceived risk, and skepticism about humanity. Vishwanath (2015) studied the impact of habitual Facebook usage on antiphishing behavior on a social media platform. Wang et al. (2017) examined the impact on antiphishing behavior of coping adaptiveness, anxiety, perceived detection efficacy, perceived susceptibility, perceived severity, and other demographic factors. Another stream of research has focused on developing effective antiphishing training programs or warning systems. For example, Kumaraguru et al. (2007a; 2007b) evaluated the effectiveness of an embedded training system to help people detect phishing. Sheng et al. (2007) and Sheng et al. (2010) tested the effectiveness of educational cartoons, computer games, and web-based training materials. Arachchilage and Love (2013) tested the effectiveness of a mobile game. Jensen et al. (2017) compared the effectiveness of mindfulness and rule-based training. Dincelli and Smith (2020) showed the advantage of a gamified security training program over traditional email-based training for reducing online self-disclosure, which is a main informational source for phishing attacks. Abbasi et al. (2021) showed that a warning system based on a phishing susceptibility prediction was more effective than traditional warning systems in reducing risky behavior during interaction with phishing websites. The study also summarized the factors that can predict individuals' phishing susceptibility: the effectiveness of antiphishing tools, individual perceptions of antiphishing tools, phishing characteristics, threat perceptions, demographic factors, and past online experience. Nguyen et al. (2021) studied the impacts of different design features of a phishing warning system on individuals' antiphishing performance. Meanwhile, the third stream of research has focused on the effectiveness of phishers' strategies to deceive people. For example, Wright et al. (2014) studied the effectiveness of influence techniques such as using authority to deceive people. Goel et al. (2017) examined the effectiveness of eliciting fear of loss and anticipation of gain. Appendix A summarizes prior studies of antiphishing behavior and performance.

As outlined in the Introduction, the conflicts between security-related goals and other work-related goals are common in individuals' daily work. Employees may focus on the goal of finishing their main task quickly while devoting little effort to their organization's security training program. Students may focus on the goal of quickly finishing their assignments while forgetting the goal of checking their incoming emails attentively to avoid phishing messages. Despite this considerable phishing literature, to the best of our knowledge, prior phishing-related research mostly focuses on antiphishing itself but doesn't take into account individuals' pursuits of other work goals. We still have limited understanding about how an individual's pursuit of other work-related goals conflicts with that person's goal pursuit of combating phishing attacks.

In addition, prior research discussed the influence of individual factors such as prior experiences, self-efficacy, motivation and emotion on antiphishing performance but ignored the influence of an antiphishing climate (i.e., individuals' perceptions of the antiphishing practices conducted by organizations where they perform their daily work). Antiphishing climate is also a factor influencing the goal striving and goal attainment of combating phishing emails in addition to individual characteristics such as self-efficacy and motivation. Individuals' goal-driven behavior and goal attainment of combating phishing emails are influenced by their antiphishing climate as well as their individual characteristics, while at the same time conflicting with their pursuits of other work-related goals. We thus believe that combining the goal-oriented framework and antiphishing climate to explain individuals' goal-driven behavior and goal attainment in combating phishing emails will provide a more comprehensive view of their antiphishing behavior and performance.

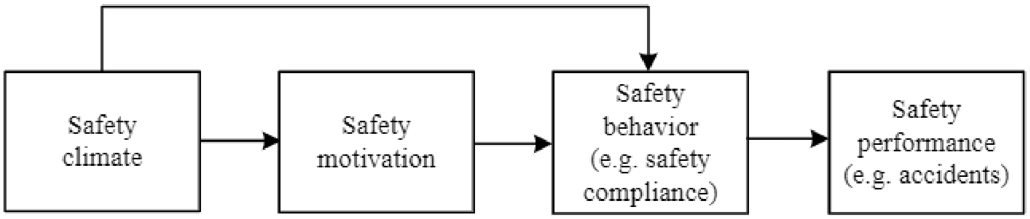
2.2. Model of Safety Performance

Research on safety performance has been mainly concerned with explaining accidents and injuries in daily work. Figure 1 shows a model of safety performance by Neal and Griffin (2006) that posits that safety climate affects safety motivation, both of which drive safety behaviors (e.g., safety compliance). These behaviors in turn determine safety performance. Safety climate is defined as an individual's perceptions of organizational values, education, and practices related to safety in daily work (Neal et al., 2000). Safety motivation is defined as an individual's needs and wants in his or her efforts to maintain safety (Neal & Griffin, 2006). Safety climate is believed to have a positive impact on safety motivation (Neal & Griffin, 2006). Both safety climate and safety motivation are known to promote safety behaviors (Neal & Griffin, 2006; Neal et al., 2000) that consist of safety compliance and safety participation. While safety compliance represents "the core activities that individuals need to carry out in order to maintain workplace safety," safety participation means behaviors that "help to develop an environment that supports safety" (Neal & Griffin, 2006, pp. 947).[2] Whereas safety motivation represents motivational forces that drive safety-relevant measures and steps, safety compliance and safety participation indicate actual endeavors. In addition, safety behaviors reduce accidents in daily work (Neal & Griffin, 2006). This safety model has been successfully applied as a powerful theoretical framework for a variety of settings, including medical (Singer et al., 2009), rail industry (Morrow et al., 2010), and chemical industry (Vinodkumar & Bhasi, 2009).

---

[2] Given the highly personal nature of antiphishing that centers on personal email accounts, such prosocial behaviors as safety participation are beyond the scope of this study.

Figure 1. Model of Safety Performance (Neal & Griffin, 2006)

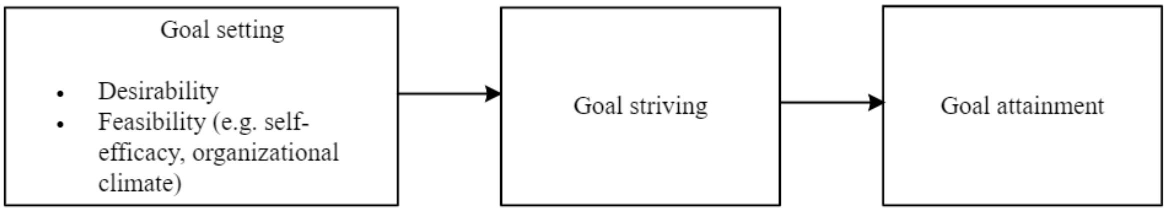

Safety climate
Safety motivation
Safety behavior (e.g. safety compliance)
Safety performance (e.g. accidents)

Although the safety model was mostly used to explain accidents and injuries in an offline work environment, it may also be extended to other potential accidents and to harm encountered when conducting online activities such as checking emails. In a traditional offline work environment, individuals want to perform their work (e.g., caring for patients) safely but still suffer from physical accidents (e.g., getting infectious diseases from patients) that can harm their physical well-being. When checking emails (e.g., emails related to their jobs, studying, or other daily tasks), individuals also desire to check their emails safely but accidentally fall for phishing emails that can damage their computers or data assets. Given that antiphishing is a form of safety-related activity, the safety model is likely to be helpful in explaining antiphishing performance. The model of safety performance specifically emphasizes the effect of a safety climate on individuals' safety motivations and behavior. It thus points out the necessity, when examining online safety-related activities such as combating phishing messages, of considering individuals' perceptions of their daily work environment as well as their perceptions of their own ability and motivation. Although it yields valuable insights into understanding antiphishing behavior, the existing model of safety performance is nevertheless limited in providing a theoretical account of some key issues. Specifically, it does not explicitly address a balance between work and safety. Although safety is one of the goals, task performance in daily work is an equally important criterion that deserves an individual's considerable attention. In summary, although the safety model was developed for physical safety, it is likely to be extended to provide a theoretical account for online safety such as antiphishing. However, as noted previously, it doesn't explicitly address the challenge of achieving safety performance while simultaneously maintaining workplace productivity.

2.3. Antiphishing as a Goal in Daily Work

Successfully combating phishing is a goal. Individuals can set a goal of combating phishing emails after learning from online resources or formal security training programs about the potential damage of phishing to their own personal information and their work environment. Therefore, to develop a theoretical explanation of antiphishing behavior, it is crucial to thoroughly account for activities such as goal setting, goal striving, and goal attainment in relation to antiphishing. Goal-oriented models have been widely used in a variety of research areas, including psychology (Sheldon & Elliot, 1999), marketing (Bagozzi & Dholakia, 1999), management (Fishbach & Choi, 2012), and information systems (Goes et al., 2016; Loock et al., 2013). Figure 2 depicts a goal-oriented framework by Bagozzi and Dholakia (1999) that describes three major activities — goal setting, goal striving, and goal attainment — involved in goal-directed behavior. First, people are said to choose a goal based on its desirability and feasibility (Gollwitzer, 1996). Whereas desirability is a motivational factor related to needs and wants, feasibility is concerned with personal characteristics (e.g., self-efficacy) as well as environmental factors (e.g., organizational climate). Second, goal striving, which involves the deployment of action plans and the initiation of action, is known to be positively associated with goal desirability and feasibility (Sheldon & Elliot, 1999). Finally, goal attainment refers to a positive consequence of goal striving.

Figure 2. Goal-Oriented Framework (Bagozzi & Dholakia, 1999)



The goal-oriented framework suggests several ways to enhance the model of safety performance as a theoretical tool for understanding antiphishing behavior. First, a safety climate corresponds well to the feasibility concept that incorporates the notion of work environment. The goal-oriented framework also implies that in addition to safety climate, self-efficacy plays an important role in determining safety performance. Moreover, safety motivation is analogous to goal desirability in the sense that both concepts represent needs and wants. Further, safety behavior corresponds well with goal striving in the sense that both deal with the implementation of action plans. It should be noted that goal striving initially requires deliberate planning and execution; however, in a highly routine environment (e.g., email checking), it can become automatic with repetition over time (Kim, 2009; Polites & Karahanna, 2012). Appropriate habitual behaviors (e.g., paying attention to phishing cues), when becoming automatic, can still lead to a high level of goal attainment (e.g., high-level of antiphishing performance). However, inappropriate habitual behaviors can hinder goal attainment. Finally, safety performance is a context-specific factor equivalent to goal attainment.

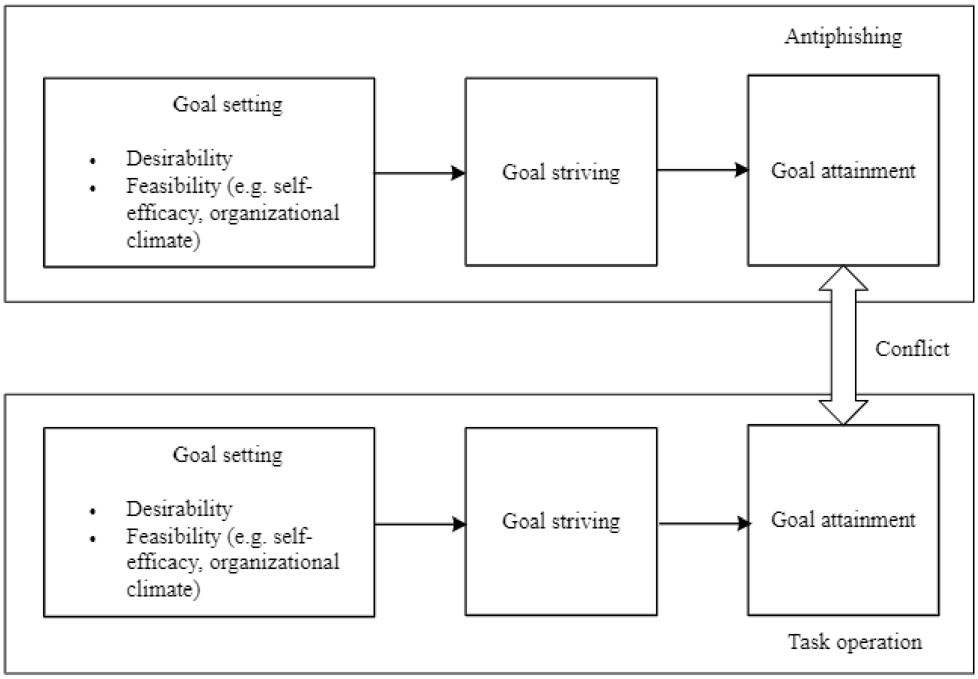
Meanwhile, goals often conflict with each other in everyday life. In such a case, goal attainment in one task could negatively affect goal attainment in other tasks (Louro et al., 2007; Vancouver et al., 2010). As Figure 3 depicts, the goal of fending off phishing emails can conflict with the pursuit of other goals (e.g., keeping up with work emails).

Specifically, a person may develop a general sense of self-efficacy to combat phishing emails when she learns from some online resources. Accordingly, she will probably make decision to fend off phishing emails when checking her email account (i.e., goal setting for antiphishing). She will then act according to the preset goal when checking emails (i.e., goal striving for antiphishing). At the same time, however, the same person may consider that checking work emails in time may facilitate her work performance. This extra goal is also not hard to achieve as long as the person checks her emails regularly. As a result, she also sets an additional goal to respond to every work email in time (i.e., goal setting for task operation). She will also act according to this goal when receiving work emails (i.e., goal striving for task operation).

A conflicting situation can arise when one goal is too salient in one’s mind compared with another (Bagozzi & Dholakia, 1999). For example, someone may only focus on a goal of checking work emails and click routinely on an email link in a work-related email without carefully checking the sender’s email address.

Thus, the goal-oriented framework sheds light on the importance of examining antiphishing in relation to other competing goals. In general, the goal-oriented framework is rich in its implications and has great potential to aid understanding of the mechanism driving antiphishing-related perceptions and performance. It is especially highly consistent with the model of safety performance, but encompasses broader contexts, including not only typical work that individuals engage in to earn salaries but also other daily tasks and duties. As a result, the goal-oriented framework can account for the potential conflict between antiphishing and other goals relating to not only typical work for salaries but also to other daily tasks such as purchasing products, studying courses, and volunteering. The goal-oriented framework can thus serve as a theoretical basis for explaining online safety-related behaviors such as antiphishing.

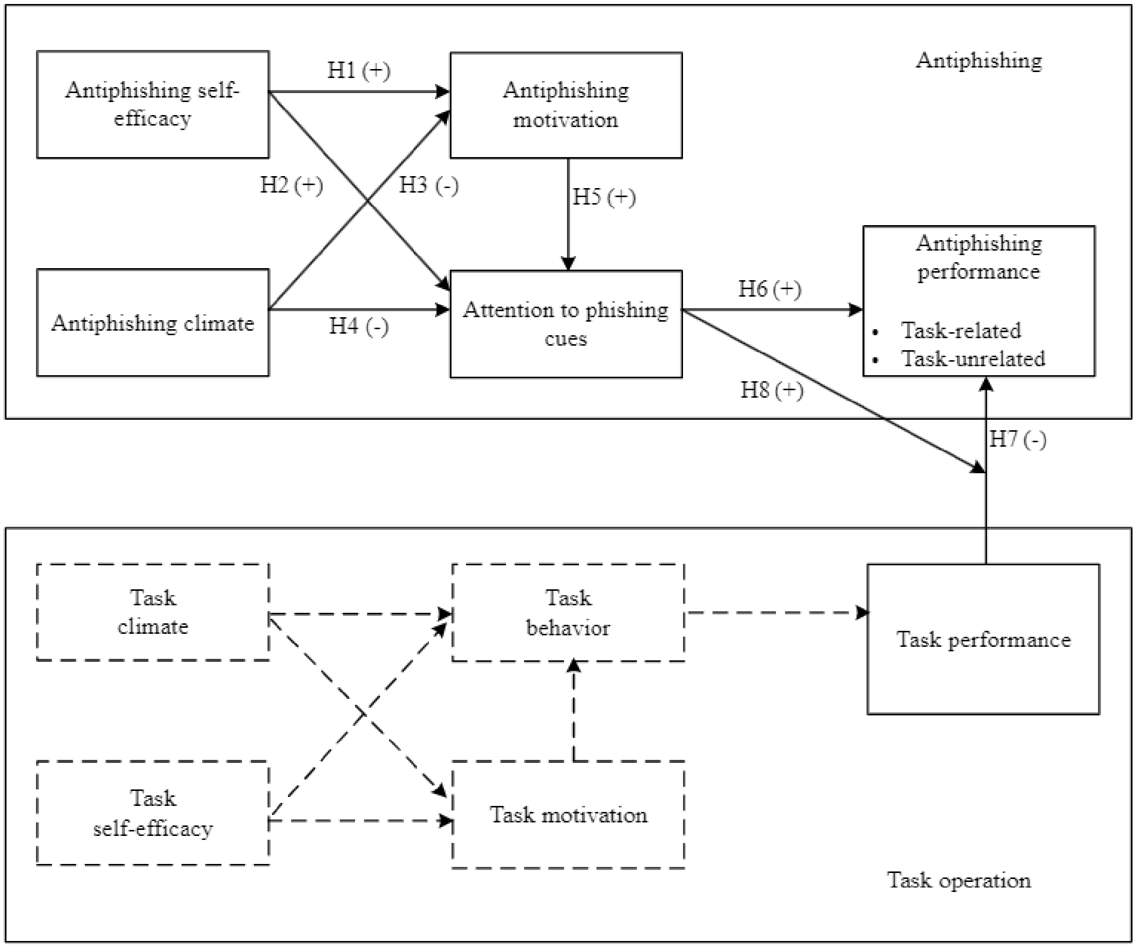
Figure 3. Conflicting Goals


Antiphishing
Goal setting
Desirability
Feasibility (e.g. self-efficacy, organizational climate)
Goal striving
Goal attainment
Conflict
Goal setting
Desirability
Feasibility (e.g. self-efficacy, organizational climate)
Goal striving
Goal attainment
Task operation

3. Research Model and Hypotheses

Figure 4 shows a proposed model of antiphishing performance, which draws on the model of safety performance complemented by the goal-oriented framework. As suggested by the goal-oriented framework, antiphishing self-efficacy is predicted to affect antiphishing motivation. Besides antiphishing self-efficacy, antiphishing climate is similarly specified as a determinant of antiphishing motivation as implied in the safety model as well as in the goal-oriented framework. Moreover, the motivation factor is expected to have a positive impact on attention to phishing cues, which is a concept that corresponds to safety behavior and specifically indicates the extent of goal striving in the context of antiphishing. We also predict that attention to phishing cues will exert a positive impact on antiphishing performance.

Figure 4. Conceptual Model


Antiphishing
Antiphishing self-efficacy
H1 (+)
Antiphishing motivation
H2 (+)
H3 (-)
H5 (+)
Antiphishing climate
H4 (-)
Attention to phishing cues
H6 (+)
Antiphishing performance
• Task-related
• Task-unrelated
H8 (+)
H7 (-)
Task climate
Task behavior
Task performance
Task self-efficacy
Task motivation
Task operation

Notes:

- Solid lines represent the foci of this study, whereas dotted lines are outside its scope.
- Control variables: perceived threat, phishing anxiety, age, gender, desktop use, laptop use, tablet use, smartphone use, email load, and prior phishing experience.

The model also indicates that individuals' perceptions, behavior, and performance concerning task operations will be similarly regulated as described previously in the case of antiphishing. Although the antecedents of task performance are outside the scope of this study, the proposed model describes the relationship between task performance and antiphishing performance. The model specifically predicts that attention to phishing cues will act as a moderator between task performance and antiphishing performance. Moreover, we also evaluated how the aforementioned relationships vary with respect to the types of phishing messages (i.e., task-related vs. task-unrelated). Each of the research hypotheses in this study is discussed in detail later in this section.

Competence is one of the most important factors in fostering human motivation (Ryan & Deci, 2000). Without the knowledge and skills required to complete a task, a person is unlikely to be truly motivated to embark on pursuing it. In IS research, self-efficacy is known to represent an IT user's competence, which affects a number of IT-related perceptions such as behavioral intention (Johnston & Warkentin, 2010), perceived avoidability (Liang & Xue, 2009), exploitive technology adaptation (Schmitz et al., 2016), and online community engagement (Ray et al., 2014).

Antiphishing self-efficacy is one's confidence about his or her capability to distinguish between phishing and legitimate emails. Research on human motivation shows that a high level of self-efficacy facilitates one's motivation to perform a behavior because the likelihood of success increases when a person possesses the relevant knowledge and skills (Ryan & Deci, 2000). Moreover, the theory of planned behavior also indicates that if a person believes he or she has the ability to perform a certain behavior, his or her intention to perform that behavior is very likely to increase (Ajzen, 1985; Bulgurcu et al., 2010; Moody et al., 2018). In the context of antiphishing, self-efficacy is shown to reduce a concern for getting phished and at the same time, enhance a rational approach to antiphishing with less emotional or maladaptive coping responses (Wang et al., 2017). Above all, we expect that antiphishing self-efficacy has a positive impact on antiphishing motivation, which refers to a commitment to keep oneself secure from phishing attacks. In fact, in safety literature, although not explicitly examined, empirical evidence has shown a high level of correlation between knowledge and motivation ($r = 0.65$) (Neal et al., 2000). Thus, we believe that if someone is confident in her ability to guard against phishing threats, she is likely to aspire to perform antiphishing behavior.

Self-efficacy is a major determinant of an individual's choice of activities, effort, and persistence (Bandura, 1977, 1997; Gist & Mitchell, 1992). Thus, we also propose that antiphishing self-efficacy will positively influence attention to phishing cues; this refers to the actual checking of one's emails with attention to their legitimacy and with security awareness (i.e., individuals' actual activities to combat phishing emails). Specifically, the rationale behind this proposition is that a person's confidence in his or her knowledge and skills has been shown to facilitate a more cognitive or rational approach, in contrast to emotional or impulsive approaches, to the evaluation of emails (Wang et al., 2017). This line of reasoning suggests that antiphishing self-efficacy helps overcome inappropriate habitual behavior when checking emails and leads to appropriate antiphishing behavior. Taken together, we hypothesize that antiphishing self-efficacy positively affects antiphishing motivation and attention to phishing cues. We state these hypotheses as follows:

H1: Antiphishing self-efficacy is positively associated with antiphishing motivation.

H2: Antiphishing self-efficacy is positively associated with attention to phishing cues.[3]

Next, antiphishing climate refers to a person's beliefs about his or her organization's commitment to antiphishing practices. The practices include not only educational programs but also specific technical safeguards against phishing attacks. Within the framework of safety performance, safety climate is theorized to positively influence safety motivation and safety behavior (Neal & Griffin, 2006; Neal et al., 2000). Specifically, based on social exchange theory, the safety model indicates that if an organization's managers care for its members' safety, members feel obligated to reciprocate for what they have received by doing as the organization asks (Neal & Griffin, 2006). Thus, the safety model predicts that safety climate is positively associated with safety motivation and with safety behavior.

---

[3] Although we hypothesized that antiphishing self-efficacy was positively associated with attention to phishing cues, the relationship between the two constructs can be more complicated than mere association. Most of the studies showed someone with higher self-efficacy than others would exert more effort and show better performance in his/her task at the between-individual level (Bandura, 1997; Stajkovic & Luthans, 1998). In contrast, at the within-person level, however, higher self-efficacy at a time could lead to lower subsequent efforts and performance (Vancouver et al., 2001; Vancouver et al., 2002; Yeo & Neal, 2006). The differential effects of self-efficacy at the within- and between-person levels can also occur in the context of antiphishing.

Nevertheless, some factors related to contextual differences lead us to expect that in the context of antiphishing, climate is negatively, instead of positively, associated with motivation and behavior. For example, safety in a traditional work environment can be achieved mostly by following prescribed guidelines and regulations. In contrast, in the case of antiphishing, hackers constantly attempt to exploit their targets with ever newer forms of deception. In this hostile online environment, individuals should maintain realistic expectations about what their organization can do to thwart novel phishing attempts. It is no longer sufficient for individuals to passively follow organizational mandates; instead, they must be flexible in adapting to the dynamically evolving landscapes of social engineering.

Much research on online behavior indicates that trust leads to risk-taking behaviors (Malhotra et al., 2004; Nicolaou & McKnight, 2006; Pavlou, 2003). It implies that if people rely heavily on an organization’s competence, honesty, and benevolence, they let down their guard against potential harm (Gefen, 2002). Moreover, research on human cybersecurity behavior also shows the evidence that in large-size organizations, people tend to assume that their organizations have done sufficient work to enhance cybersecurity and their data asset has been taken good care of regardless of their own efforts (Chowdhury et al., 2018). These findings point out that members of an organization who have a high level (or even an inappropriately high level) of confidence in its protection of information are likely to become negligent in exercising their own due diligence. Thus, they are less likely to attribute importance to proactive antiphishing and pay less attention to phishing cues. Thus, we predict that because of the dynamically changing and increasingly sophisticated online environment, the perceived climate concerning antiphishing will have negative impact on individuals’ antiphishing motivation and antiphishing behavior.[4]

H3: Antiphishing climate is negatively associated with antiphishing motivation.

H4: Antiphishing climate is negatively associated with attention to phishing cues.

---

[4] We built H3 and H4 while assuming that individuals’ main responsibilities are not necessarily security-related tasks. The hypotheses will not be applicable to individuals who are involved mainly in the security sector; their roles in the organizations may more strongly influence antiphishing motivation and attention to phishing cues.

Goal pursuit is expected to be stronger when goal desire is higher (Bagozzi, 2007; Bagozzi et al., 2003). Therefore, the greater the desire to combat phishing, the stronger a person's mindful checking of phishing cues. However, with repeated exposure to legitimate emails, such caution tends to decline and eventually enters the realm of routinization (Vishwanath, 2015). Routinized practices often lead to desirable outcomes in numerous circumstances (e.g., learning, physical exercise). Thus, at least in such cases, high motivation could lead to automatic behavior, which eventually leads to high performance. However, antiphishing is unique in that hackers can exploit a lack of attention. Because of the potential risks constantly present in this online environment, people need to resist a natural tendency to become mindless in identifying phishing cues (Brown et al., 2007).

In a field study of antiphishing performance, Jensen et al. (2017) found that persons who were taught the importance of being mindful of peculiarity and irregularity in email messages could identify phishing attempts better than others who lacked such a mindfulness training. These findings imply that to avoid falling a victim to phishing, people should understand the importance of avoiding habitual or automatized reactions and instead focus on situational awareness. The discussion mentioned previously suggests that those who are motivated to guard against phishing threats are likelier to overcome routinized or overlearned patterns of checking emails and remain attentive to phishing cues. Thus, we propose that antiphishing motivation will have a positive impact on attention to phishing cues.

H5: Antiphishing motivation is positively associated with attention to phishing cues.

According to the model of safety performance, safety behavior is hypothesized to increase safety performance. Similarly, we propose that attention to phishing cues, which is a form of antiphishing behavior, will have a positive impact on antiphishing performance. The task of differentiating between legitimate and phishing emails requires conscious attention; accordingly, a lack of attention to phishing cues will make people vulnerable to phishing attempts. In a study of Facebook use, Vishwanath (2015) showed that habitual Facebook users respond to phishing messages more often than non-habitual users. These findings suggest that people tend to get phished more when they respond to messages habitually with little conscious effort. That is, a high level of antiphishing performance calls for a high level of attention in inspecting messages. Accordingly, we hypothesize that attention to phishing cues will have a positive impact on antiphishing performance.

H6: Attention to phishing cues is positively associated with antiphishing performance.

Given that attentional capacity is limited, competing cognitive demands arising from conflicting activities are unlikely to be fully met. Accordingly, performance improvement for one goal is likely to result in performance deterioration for another (Strayer & Johnston, 2001; Strayer et al., 2003). More specifically, when a person is conscientious in timely responding to work emails, she is likelier to click on legitimate emails. However, because of the extra attention paid to completing this task, she may ignore suspicious information indicative of a phishing attempt. Thus, we predict that those who demonstrate better task performance in keeping up with work emails (i.e., those who click more on legitimate work emails) are expected to click more on phishing emails.

In addition, we also propose that the effect of task performance on antiphishing performance will be more evident for task-related phishing messages than for phishing messages unrelated to tasks. As noted earlier, focusing on a main task requires considerable mental effort and leaves little capacity for a secondary, and sometimes conflicting, task (Vancouver et al., 2010). This is especially so when people are eager to pursue a focal goal and consequently tend to become somewhat inattentive to a secondary goal (Louro et al., 2007). In our case, for example, if people strive to accomplish a task such as timely management of work emails, they are likelier to overlook suspicious cues in task-related phishing emails because of their exclusive concentration on an impending task. Meanwhile, when phishing emails are not directly related to a task operation in daily work (e.g., customer rewards, personal finance), the conflict between the multiple goals (i.e., task operation and antiphishing) is less severe than in the previous case of task-related phishing messages. Thus, the impact of task performance on antiphishing performance is likely to be stronger for task-related phishing messages than for task-unrelated phishing messages.

H7: Task performance is negatively associated with antiphishing performance; specifically, the relationship between task performance and antiphishing performance is stronger for task-related phishing messages than for other phishing messages.

More focus on checking work emails may increase the possibility of succumbing for phishing messages, especially when the phishing message is related to work. Meanwhile, in the case of people more attentive to phishing cues, they can be more effective at detecting whether a message is a legitimate work email or a work-related phishing scam because of their cautiousness. Thus, we also predict that the negative impact of task performance on antiphishing performance will decrease with an increase in attention to phishing cues. As noted earlier, the negative impact of task performance on antiphishing performance results from the pursuit of conflicting goals. Diverting mental effort to one area (e.g., task operation) leads to reduced effort in another area (e.g., antiphishing), especially when the primary and secondary goals conflict. Nevertheless, it would be still possible to increase the pool of attentional resources for the two competing goals in question (Watson & Strayer, 2010). Even if a person is highly attentive to task operation and achieves a high level of task performance, she may still be able to expend a high level of attention to antiphishing when she feels such vigilance is warranted. Thus, increasing attention to antiphishing may mitigate the conflicting relationship previously mentioned. In such a case, the negative influence of task performance on antiphishing performance is likely to decrease because the extra attention set aside for antiphishing makes it possible to overcome the harmful effect of focusing too much on task performance. Moreover, because the negative impact of task performance on antiphishing performance is hypothesized to be stronger for task-related phishing emails, the moderating effect of attention is likely to be more evident for task-related phishing emails. Thus, we hypothesize that especially for task-related phishing messages, the negative effect of task performance on antiphishing performance will decrease as attention to phishing cues increases.

H8: The relationship between task performance and antiphishing performance decreases with the increase in attention to phishing cues; specifically, the moderating effect is stronger for task-related phishing messages than for other phishing messages.

The proposed model also includes a number of control variables believed to be significant in regulating antiphishing-related perceptions and behavior. For example, Vishwanath et al. (2011) suggested that email load can affect antiphishing behavior. Age, gender, and previous phishing experience have also been commonly used as control variables in related studies (Kumaraguru et al., 2007a; Wang et al., 2017). Wang et al.'s (2017) model also suggested three antecedents of antiphishing behavior, namely, perceived threat, antiphishing self-efficacy, and phishing anxiety. Because antiphishing self-efficacy has already been included in the research model, we have treated as control variables both perceived threat — defined as cognitive beliefs about the risk of getting phished — and phishing anxiety — defined as emotional feelings arising from the risk of getting phished. Moreover, evidence also suggests that the different devices used may also affect security-related behaviors (Thompson et al., 2017). Thus, we also controlled in the study for the frequency of using different devices to check emails, namely, use of desktops, laptops, tablets, and smartphones.

4. Research Method

4.1. Sample and Data Collection

To test our hypotheses, we conducted a field experiment with college students enrolled in a general business course in a public university in the United States. The experiment was approved by the Institutional Review Board.[5] We chose college students as subjects because we considered the university as an organization in which students pursue a goal of keeping up with their legitimate work emails (i.e., emails related to their course work) as well as one of thwarting phishing emails because they were also likely to encounter both task-related and task-unrelated phishing emails. We sent four legitimate work emails and four phishing emails alternately to these students over five weeks and recorded their clicking behaviors.[6] All the subjects were sent the same email at the same time and received the eight emails in the same order. The students were told nothing about the experiment during the five weeks to keep their email clicking behaviors as natural as possible. In the sixth week, we conducted a survey to collect data about antiphishing self-efficacy, antiphishing climate, antiphishing motivation, attention to phishing cues, and other control variables. We provided additional course credits to the students for participating in the survey. Our original sample consisted of 357 subjects. However, 20 potential subjects did not participate in the survey. Thus, we received 337 valid responses to the survey. After the survey, the participants were informed about the previous field experiment. Among the 337 participants, 49% were female, and their average age was 20.38 years.

---

[5] The Institutional Review Board approved the experiment, which has reference ID 2017-0914.

[6] We used Barracuda Phishline (https://www.barracuda.com/products/phishline) to send phishing emails in this experiment.

Appendix B lists all the legitimate and phishing emails used in the experiment. We embedded links within the emails to track students' clicks. Specifically, all the legitimate emails were about class materials, introductions of the instructor and TAs, and course topics. If a student clicked on the link in a legitimate email, he or she was identified as successfully performing the task of checking this legitimate email. Otherwise, he or she was recognized as not successfully performing this task. Upon clicking on the link, the student would be directed to the university's course website to see more specific information, and the course website recorded his or her clicking behavior. The students could only access the detailed course materials by clicking on the links in the legitimate emails. For phishing emails, the first of the phishing emails we designed mimicked an email from the course; the second imitated a gift card offer from the business school; the third simulated a security breach alert from Google; and the last mimicked a course notification from the university's learning management system. Therefore, the second and third phishing emails were task-unrelated (i.e., unrelated to keeping up with course emails). The email addresses of all the phishing emails were designed to sound fake (e.g., administrator@remoteemail.net). Moreover, if a student hovered his or her mouse on the link in a phishing email, he or she saw a suspicious URL (e.g., remoteemail.net). As in the method used by Goel et al. (2017), when a student clicked on the link, an error message popped up, and the software used for sending the phishing emails identified the student as being successfully phished. Otherwise, he or she was recognized as avoiding the phishing attack. We tracked their clicking behavior for each email for four days after the email was sent. Similar research has shown that within eight hours after an email is sent, nearly 90% of the people who will eventually click on the link have already done so (Kumaraguru et al., 2009). It is reasonable to believe that people seldom return to emails received four days earlier.

4.2. Measures

Table 1 contains descriptions of how we measured all the research variables. We adapted most of the survey items from past safety literature and IS literature to fit our research context. We measured age and gender with single-item measures. We asked our participants to report the frequency with which they used these four types of devices, desktop, laptop, tablet, and smartphone use, to check for emails. We also measured email load with the same method as Vishwanath et al. (2011). Prior phishing experience was measured with a single-item scale. Perceived threat consisted of one item for perceived susceptibility and one item for perceived severity. We adopted the items used by Wang et al. (2017). We then multiplied the two items to form a single indicator for perceived threat. We measured phishing anxiety according to the method of Wang et al. (2017).

Items for antiphishing climate and antiphishing motivation were adapted from Neal and Griffin (2006). We adapted the items by changing their offline context to the context of antiphishing. Items for antiphishing self-efficacy were adapted from Kankanhalli et al. (2005) and Venkatesh et al. (2003). We adapted the Mindful Attention Awareness Scale (MAAS) from MacKillop and Anderson (2007) into a phishing context to measure attention to phishing cues. We reversed the coding of attention to phishing cues in the data analysis so that in our results a higher score represented more attention. We first focused on general antiphishing performance as measured by a person's total times of avoiding phishing attacks across the four phishing attacks. We then focused on more specific antiphishing performance (whether a student clicked on the link) for each phishing attack. We coded it as 1 if a person did not click on the link in a phishing email, and 0 if he or she did. We measured task performance as the total number of clicks on the legitimate course emails.

Table 1. Measurement Items

Constructs	Item	Source
Control Variables		
Age (AGE)	Age in years	Self-developed
Gender (GEN)	1=Female; 0=Male	Self-developed
Desktop use (DSK)	Average frequency to use desktop to check the university email account	Self-developed
Laptop use (LAP)	Average frequency to use tablet to check the university email account	Self-developed
Smartphone use (PHN)	Average frequency to use smartphone to check the university email account	Self-developed
Email load (LOAD)	Number of emails received per day on average within the university's email account	Vishwanath et al. (2011)
Prior phishing experience (EXP)	How often have you been the victim of what you think was an attempt to phish you? 5-point frequency rating (1 - Never, 2 - Very rarely, 3 - Rarely, 4 - Occasionally, 5 - Very frequently)	Self-developed
Perceived threat (THR)[a]	THR1. It is very likely that I will become a victim of phishing attacks. THR2. It would be a significant loss if I fall a victim to a phishing attack.	Wang et al. (2017)
Phishing anxiety (ANX)	ANX1. When I think about being phished, I feel nervous. ANX2. When I think about being phished, I get upset.	Wang et al. (2017)
Research Variables		
Antiphishing climate (CLM)	CLM1. Filtering phishing emails is given a high priority by the university. CLM2. The university cares about education on avoiding phishing scams.	Neal & Griffin (2006)
Antiphishing self-efficacy (EFF)	EFF1. I consider myself familiar with phishing scams. EFF2. I consider myself informed about how to avoid phishing scams. EFF3. I consider myself knowledgeable about recognizing phishing scams.	Kankanhalli et al. (2005) Venkatesh et al. (2003)
Antiphishing motivation (MOV)	MOV1. I feel that it is worthwhile to put in effort to avoid phishing scams. MOV2. I feel that it is important to avoid phishing scams at all times. MOV3. I believe that it is important to reduce the risk of being phished.	Neal & Griffin (2006)
Attention to phishing cues (ATN)[b]	Please indicate the frequency of the experience below when you check emails. ATN1. I click on links or open attachments in emails automatically, without being aware of the potential risk of being phished. ATN2. I find myself click on links or open attachments in emails without paying attention to the legitimacy of them. ATN3. "Just do it" describes the way I respond to requests in emails.	MacKillop & Anderson (2007)
Antiphishing performance (AP)	We coded it as 1 if a person didn't click on a phishing email, and 0 if he or she clicked on it. Antiphishing performance was measured by a person's aggregated score of the four phishing attacks, and also the clicking behavior for each phishing attack.	Measured objectively

Task performance (TP)	We coded it as 1 if a person clicked on a legitimate work email, and 0 if he or she didn't click on it. Task performance was measured by a person's aggregated score of the four legitimate emails.	Measured objectively

Notes:

- Items about frequency were measured with a 5-point frequency scale (1 - Never, 2 - Rarely, 3 - Sometimes, 4 - Very often, 5 - Always); other items were measured with a 5-point Likert scale (1 - Strongly Disagree, 2 - Disagree, 3 - Neither Agree nor Disagree, 4 - Agree, 5 - Strongly Agree) if not specified.
- [a] The scores of the two items were multiplied to form a single indicator for perceived threat.
- [b] The score was reversed coded as 1 - Always to 5 - Never.

5. Analyses and Results

5.1. Data Analyses

This section will first conduct a preliminary analysis to see the general click rates on our phishing and legitimate emails. We will then evaluate the psychometric properties of our measurement scales by conducting exploratory and confirmatory factor analysis.

To test our hypotheses, we will use SEM to test the proposed relationships between the latent constructs (i.e., antiphishing self-efficacy, antiphishing climate, antiphishing motivation, and attention to phishing cues) based on several model fit statistics and the path coefficients. We will also initially use SEM to analyze the relationships between attention, task performance, and the aggregated antiphishing performance regardless of email characteristics.

Moreover, especially for testing the relationships between task performance and antiphishing performance (H7 and H8), we will use generalized estimating equations to better accommodate the hierarchical data structure in which repeated, dichotomous outcomes (i.e., click or not click on an email) were nested within individuals. In this way, we can also analyze how the relationships may differ according to email characteristics.

5.2. Preliminary Results

At the beginning, a relatively large proportion of participants clicked on the phishing emails. Forty-two percent of the participants clicked on the first phishing email, and 35% clicked on the second one. At a later phase of the study, the click rates dropped slightly, to 12% for the third phishing email and 25% for the fourth one. There were no significant differences between male and female participants.

For legitimate emails, the clicking rates remained relatively stable (18%, 25%, and 17% for the first, third, and the fourth email, respectively) except for the second legitimate email (6%). There were still no significant differences between male and female participants. Table 2 shows the click rates for each email by gender.

Table 2. Clicking Rates by Gender

GENDER	LEG1	PH1	LEG2	PH2	LEG3	PH3	LEG4	PH4
Male	16%	44%	5%	33%	24%	13%	17%	25%
Female	19%	41%	6%	38%	27%	11%	16%	25%
Average	18%	42%	6%	35%	25%	12%	17%	25%

Notes:

- The order of the email from left to right indicates the real order in which they were sent.
- LEG = legitimate email; PH = phishing email.

5.3. Exploratory Factor Analysis

Our key research constructs (i.e., antiphishing climate, antiphishing self-efficacy, antiphishing motivation, and attention to phishing cues) were adapted from prior literature unrelated to phishing research. We revised the wordings of the prior items to fit our phishing context. To explore the factor structure of those survey items, we first conducted exploratory factor analysis using a principal axis factoring with promax rotation using SPSS 25. The Kaiser-Meyer-Olkin measure was 0.74, indicating that the data was relatively suitable for factor analysis. The factor analysis revealed a four-factor solution based on the criterion of eigenvalues greater than one and the inspection of the scree plot (Brown & Ryan, 2003). Specifically, the first factor consisted of three items of attention to phishing cues and accounted for 32.53% of the variance with an eigenvalue of 3.58; the second factor consisted of the three items of antiphishing self-efficacy and accounted for 14.84% of the variance with an eigenvalue of 1.63; the third factor consisted of three items of antiphishing motivation and accounted for 14.25% of the variance with an eigenvalue of 1.57; the fourth factor consisted of two items of antiphishing climate and accounted for 9.23% of the variance with an eigenvalue of 1.02. Table 3 shows the factor loadings and cross-loadings for the items.

5.4. Confirmatory Factor Analysis

To further evaluate the psychometric properties of our scales, we did a confirmatory factor analysis encompassing all 16 variables (both research variables and control variables) in our model. The original model fit the data well: $\chi2$ (143) = 147.66 (p = 0.378), RMSEA = 0.010, GFI = 0.96, CFI = 1.00, NNFI =

Table 3. Item Loadings for Key Research Constructs

Construct	Item	ATN	EFF	MOV	CLM
ATN	ATN_1	**0.87**	0.00	0.03	0.00
	ATN_2	**0.97**	0.01	-0.07	0.01
	ATN_3	**0.73**	-0.01	0.05	-0.02
EFF	EFF_1	-0.04	**0.81**	0.03	-0.06
	EFF_2	0.00	**0.87**	-0.02	0.02
	EFF_3	0.03	**0.91**	0.00	0.03
MOV	MOV_1	-0.02	0.02	**0.73**	0.02
	MOV_2	0.01	-0.01	**0.90**	-0.02
	MOV_3	0.02	0.00	**0.90**	0.01
CLM	CLM_1	-0.02	-0.01	-0.02	**0.80**
	CLM_2	0.02	0.00	0.03	**0.71**

Notes:

- CLM = antiphishing climate; EFF = antiphishing self-efficacy; MOV = antiphishing motivation; ATN = attention to phishing cues.

0.99. The values of Cronbach's alpha were all above 0.70, indicating the internal consistency of our items. We evaluated reliability through composite reliability (CR) and average variance extracted (AVE). All CRs exceeded the cutoff value of 0.70, and AVEs were above 0.50 (Bagozzi & Yi, 1988; Fornell & Larcker, 1981). Accordingly, the multi-item scales demonstrated satisfactory reliabilities. We used factor loadings to check convergent validity and found that all factor loadings exceeded 0.60 (Chin et al., 1997). Finally, we compared two models for each pair of factors: one in which the pair was allowed to correlate freely and the other in which the pair was set to correlate perfectly (Segars, 1997). The chi-square difference tests were all significant and provided evidence for discriminant validity. Table 4 shows the means, standard deviations, and correlations between all the variables. Table 5 shows Cronbach's alpha, composite reliabilities, average variance extracted of the measures for our multi-item constructs.

Common method variance has been a concern in behavioral research. As stated earlier in the research method section, we attempted to reduce such bias by measuring our independent and dependent variables using different methods at different times (Sharma et al., 2009).

Table 4. Means, Standard Deviations and Correlations of Measurement Scales

Construct	ME	SD	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
			Correlation Matrix															
1. AGE	20.38	1.53	-															
2. GEN	0.49	0.50	-0.07	-														
3. DSK	1.82	1.06	0.02	0.04	-													
4. LAP	4.17	0.78	-0.01	0.20**	0.09	-												
5. TAB	1.33	0.74	-0.02	0.03	0.35**	0.08	-											
6. PHN	4.17	0.84	-0.05	0.11*	0.00	-0.09	0.18**	-										
7. LOAD	19.36	15.62	-0.01	-0.02	0.09	0.07	-0.04	0.01	-									
8. EXP	3.08	1.07	-0.03	0.13*	0.12*	0.02	0.11*	0.09	0.07	-								
9. THR	11.81	5.72	0.00	0.23**	0.05	0.04	0.10	0.19**	0.07	0.24**	-							
10. ANX	3.68	1.04	-0.02	0.27**	0.08	0.11*	0.12*	0.09	0.09	0.20**	0.36**	**0.84**						
11. CLM	3.12	0.97	0.10	-0.03	0.05	-0.02	0.04	0.03	-0.03	-0.13*	-0.08	-0.04	**0.75**					
12. EFF	3.74	0.94	0.03	-0.10	0.08	-0.06	-0.11*	0.03	-0.07	0.00	-0.23**	-0.13*	0.11*	**0.86**				
13. MOV	4.57	0.56	0.03	0.10	0.00	0.08	-0.12*	0.10	0.06	0.09	0.14*	0.31**	0.06	0.26**	**0.85**			
14. ATN	3.82	0.86	-0.01	0.05	-0.09	-0.08	-0.17**	-0.05	-0.09	-0.07	-0.19**	-0.14*	-0.06	0.31**	0.25**	**0.84**		
15. AP	2.85	1.00	0.04	0.00	-0.03	-0.09	-0.12*	-0.13*	0.07	-0.18**	-0.12*	-0.10	0.00	0.07	-0.02	0.23**	-	
16. TP	0.65	0.84	-0.02	0.03	0.06	0.06	0.07	-0.01	-0.05	0.03	0.08	0.05	-0.04	0.07	0.14**	-0.04	-0.26**	-

Notes:

- N = 337
- ME = mean; SD = standard deviation.
- GEN = gender; DSK = desktop use; LAP = laptop use; TAB = tablet use; PHN = smartphone use; LOAD = email load; EXP = prior phishing experience; THR = perceived threat; ANX = phishing anxiety; CLM = antiphishing climate; EFF = antiphishing self-efficacy; MOV = antiphishing motivation; ATN = attention to phishing cues; AP = antiphishing performance; TP = task performance.
- The diagonal elements in bold are the square root of AVEs.
- $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$ (two-tailed).

Table 5. Statistics for Internal Consistency and Reliability

Construct	Cronbach's Alpha (α)	Composite Reliability	AVE
ANX*	0.81	0.82	0.70
CLM	0.72	0.72	0.57
EFF	0.90	0.90	0.74
MOV	0.88	0.88	0.72
ATN	0.89	0.88	0.70

Notes:

- ANX = phishing anxiety; CLM = antiphishing climate; EFF = antiphishing self-efficacy; MOV = antiphishing motivation; ATN = attention to phishing cues.
- * The variable was used as a control variable in the subsequent analysis.

5.5. Structural Equation Models

Our proposed model consisted of a series of structural relations among latent constructs. Therefore, we first used structural equation modeling (SEM) with LISREL 8 to test our research hypotheses. The proposed model represents our hypotheses: antiphishing motivation is influenced by both individuals' antiphishing self-efficacy and the antiphishing climate, as shown in Equation 1; attention to phishing cues is influenced by the antiphishing climate, antiphishing self-efficacy and antiphishing motivation as shown in Equation 2; further, antiphishing performance is influenced by attention to phishing cues, task performance, and their interaction, as shown in Equation 3.

The proposed model is shown as follows:

$$\eta_1 = \gamma_{11}\xi_1 + \gamma_{12}\xi_2 + \sum_{k=5}^{14}\gamma_{1k}\xi_k + \zeta_1$$

(1)

$$\eta_2 = \gamma_{21}\xi_1 + \gamma_{22}\xi_2 + \beta_{21}\eta_1 + \sum_{k=5}^{14}\gamma_{2k}\xi_k + \zeta_2$$

(2)

$$\eta_3 = \gamma_{33}\xi_3 + \gamma_{34}\xi_4 + \beta_{32}\eta_2 + \sum_{k=5}^{14}\gamma_{3k}\xi_k + \zeta_3$$

(3)

Notes:

- η_i – endogenous variables: η_1 = antiphishing motivation; η_2 = attention to phishing cues; η_3 = antiphishing performance.

- ξ_i – exogenous variables: ξ_1 = antiphishing self-efficacy; ξ_2 = antiphishing climate; ξ_3 = task performance; ξ_4 = interaction term between attention to phishing cues and task performance; ξ_k = 10 control variables.
- γ_{ik} – path coefficients for exogenous variables.
- β_{ik} – path coefficients for endogenous variables.
- ζ_i – residual error variances.

We also developed a baseline model that includes only control variables. In the baseline model, we entered the 10 control variables into our model for each of the endogenous variables: antiphishing motivation, attention to phishing cues, and antiphishing performance. Table 6 displays our results. Model 1(the baseline model) fit the data relatively well: $\chi 2$ (155) = 290.51 ($p < 0.001$), RMSEA = 0.051, GFI = 0.93, CFI = 0.96, NNFI = 0.92. However, Model 2, our proposed model, fit the data much better: $\chi 2$ (158) = 164.44 ($p = 0.347$), RMSEA = 0.011, GFI = 0.96, CFI = 1.00, NNFI = 0.99. This model explained 28% of the variance for antiphishing motivation, 25% for attention to phishing cues, and 18% for antiphishing performance. Figure 5 depicts the results of the proposed model, including path estimates and their statistical significance.

As hypothesized, antiphishing self-efficacy was significantly and positively associated with antiphishing motivation (standardized coefficient = 0.35, $p < 0.001$, one-tailed), supporting H1. Antiphishing self-efficacy was also significantly and positively associated with attention to phishing cues (standardized coefficient = 0.24, $p < 0.001$, one-tailed), supporting H2. However, antiphishing climate was not significantly negatively associated with antiphishing motivation (standardized coefficient = 0.05, $p = 0.21$); thus, H3 was not supported. In contrast, antiphishing climate was significantly and negatively associated with attention to phishing cues (standardized coefficient = -0.14, $p < 0.05$, one-tailed), supporting H4. Meanwhile, antiphishing motivation was significantly and positively associated with attention to phishing cues (standardized coefficient = 0.27, $p < 0.001$, one-tailed), supporting H5.

Regarding antiphishing performance, attention to phishing cues was significantly and positively associated with general antiphishing performance without discriminating between task-related and task-unrelated phishing emails (standardized coefficient = 0.22, $p < 0.001$, one-tailed), supporting H6. We then tested the relationship between task performance and antiphishing performance. We found that task performance had a negative impact on antiphishing performance in general (standardized coefficient = -0.24, $p < 0.001$, one-tailed), supporting the first part of H7. In addition, the relationship between task performance and general antiphishing performance was moderated by attention to phishing cues (standardized coefficient = 0.08, $p < 0.05$, one-tailed), supporting the first part of H8. In this SEM analysis, as discussed previously, antiphishing performance was measured by the total number of phishing emails avoided across the four phishing attacks. Accordingly, we did not take into account the differences among phishing emails nor did we discriminate between antiphishing performance for task-related and

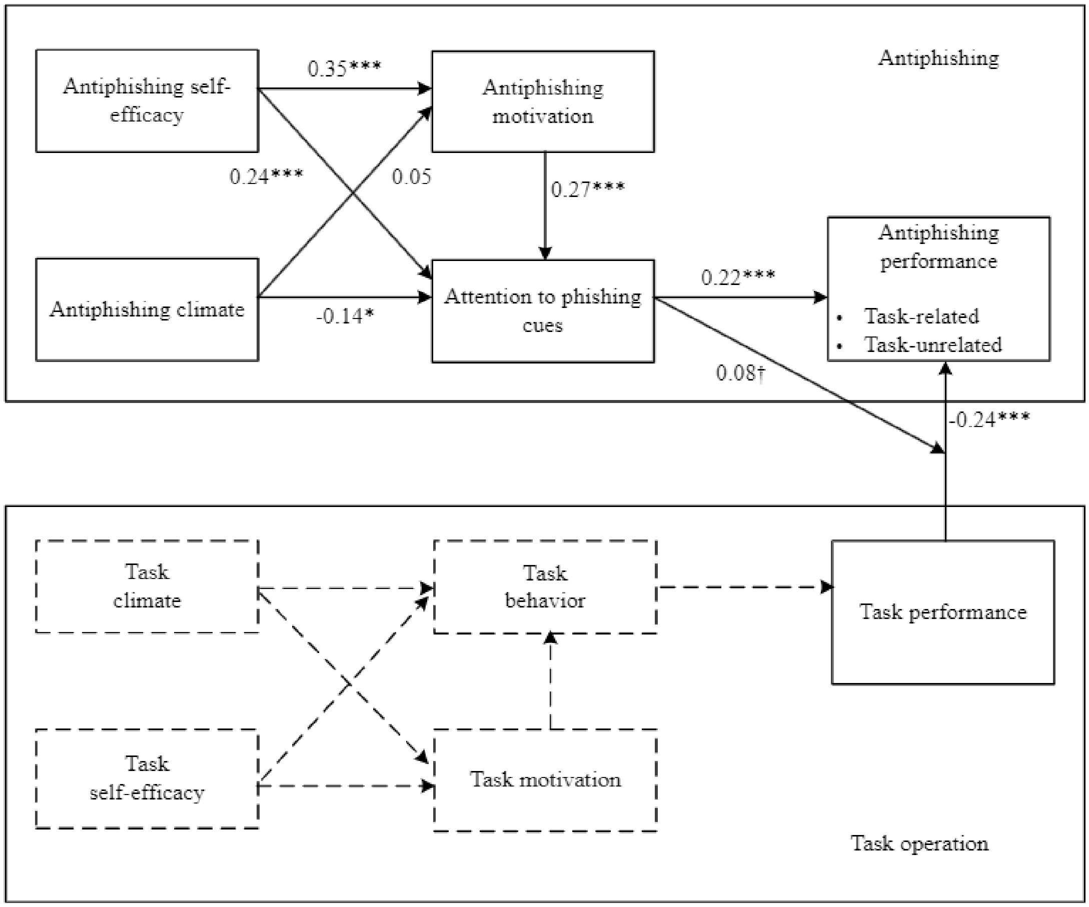
Table 6. Results of Structural Equation Modeling

Effects	Model 1			Model 2		
Causes	MOV	ATN	AP	MOV	ATN	AP
AGE	0.04	0.00	0.03	0.03	0.01	0.02
GEN	0.00	0.14*	0.09	0.01	0.16**	0.06
DSK	0.03	-0.01	0.01	-0.03	-0.05	0.02
LAP	0.06	-0.06	-0.11*	0.07	-0.08	-0.08
TAB	-0.22***	-0.13*	-0.06	-0.17**	-0.03	-0.02
PHN	0.11*	0.00	-0.11*	0.07	-0.05	-0.12*
LOAD	0.01	-0.05	0.09†	0.04	-0.04	0.09†
EXP	0.05	-0.02	-0.15**	0.03	-0.07	-0.15**
THR	0.00	-0.18**	-0.07	0.08	-0.14*	-0.02
ANX	0.31***	-0.07	-0.04	0.35***	-0.15*	-0.02
CLM				0.05	-0.14*	
EFF				0.35***	0.24***	
MOV					0.27***	
ATN						0.22***
TP						-0.24***
TP*ATN						0.08†
SMC	0.15	0.09	0.08	0.28	0.25	0.18
Model Fit						
χ^2		290.51			164.44	
d.f.		155			158	
RMSEA		0.051			0.011	
GFI		0.93			0.96	
CFI		0.96			1.00	
NNFI		0.92			0.99	

Notes:

- GEN = gender; DSK = desktop use; LAP = laptop use; TAB = tablet use; PHN = smartphone use; LOAD = email load; EXP = prior phishing experience; THR = perceived threat; ANX = phishing anxiety; CLM = antiphishing climate; EFF = antiphishing self-efficacy; MOV = antiphishing motivation; ATN = attention to phishing cues; AP = antiphishing performance; TP = task performance.
- $^{\dagger} p < 0.10$, $^{*} p < 0.05$, $^{**} p < 0.01$, $^{***} p < 0.001$ (two-tailed).

Figure 5. Results of Structural Equation Modeling


Antiphishing self-efficacy
0.35***
Antiphishing motivation
Antiphishing
0.24***
0.05
0.27***
Antiphishing performance
• Task-related
• Task-unrelated
Antiphishing climate
-0.14*
Attention to phishing cues
0.22***
0.08†
-0.24***
Task climate
Task behavior
Task performance
Task self-efficacy
Task motivation
Task operation

Notes:

- Solid lines represent the foci of this study, whereas dotted lines are outside its scope.
- Control variables: age, gender, desktop use, laptop use, tablet use, smartphone use, email load, prior phishing experience, perceived threat, and phishing anxiety.
- $^{\dagger} p < 0.10$, $^{*} p < 0.05$, $^{**} p < 0.01$, $^{***} p < 0.001$ (two-tailed).

task-unrelated phishing emails. The specific impact of task performance, and its interaction with attention to phishing cues on antiphishing performance for the two types of phishing, needed further confirmation. Nevertheless, the results based on SEM provided strong support for our proposed theoretical framework and research hypotheses.

In addition to scrutinizing the relationships among research variables, we also examined the impact of control variables on endogenous variables. First, we found that phishing anxiety had a positive relationship with antiphishing motivation. In addition, participants who frequently used their smartphones to check emails were found to show a lower level of antiphishing performance. This is probably because smartphones do not show the URL addresses associated with the hyperlinks in phishing emails. Moreover, participants with more prior phishing experience demonstrated a lower level of antiphishing performance, suggesting that people phished once tend to be phished again.

5.6. Generalized Estimating Equations

The SEM-based results supported the hypothesized relationships between task performance, the interaction between task performance and attention, and antiphishing performance. However, the SEM analysis was done in a general way with no consideration of the characteristics of the phishing emails (i.e., task-related or task-unrelated). Therefore, to better evaluate the hypotheses on the relationships between task performance, the interaction between task performance and attention, and antiphishing performance, we used generalized estimating equations (GEE) with SPSS 25 to test our hypotheses (Heck et al., 2012). GEE accommodate logistic regression for time series data while allowing the repeated observations across time to correlate with each other. It is suitable for our hierarchical data structure in which repeated, dichotomous outcomes (i.e., click or not click) were nested within individuals.

We split the data set by the email category — that is, task-related and task-unrelated phishing emails — and analyzed the subsets separately. One subset included clicking behavior for two task-related phishing emails (Phishing 1 and 4); the other included clicking behaviors for two task-unrelated phishing emails (Phishing 2 and 3). For both task-related and task-unrelated phishing emails, our Model 1 included only the Level 1 variable. Model 2 included both Level 1 and Level 2 variables. Our Level 1 variable was the time variable representing waves of phishing attacks. We coded the time variable as a dummy variable to accommodate the difference between the two phishing attacks in each subset. Our Level 2 variables were the individual level variables, including the 10 control variables as well as attention to phishing cues and task performance.

Our Model 1, including only Level 1 variables, was as follows:

$$\ln\left(\frac{\pi_{ti}}{1-\pi_{ti}}\right) = \beta_0 + \beta_1 TIME_{ti}$$

(4)

We then added the Level 2 variables as attention to phishing cues, task performance, their interaction, and the other control variables into the model. Our Model 2 was specified as follows:

$$\ln\left(\frac{\pi_{ti}}{1-\pi_{ti}}\right) = \beta_0 + \beta_1 TIME_{ti} + \beta_2 ATN_i + \beta_3 TP_i + \beta_4 TP_i * ATN_i + \sum_{k=5}^{14} \beta_k CONTROL_{ki}$$

(5)

Notes:

- π_{ti} = the probability of not clicking on a phishing email of subject i at time t; $TIME_{ti}$ = a time variable as the first phishing coded as 0 and the second phishing coded as 1 for the two task-related phishing emails and the two task-unrelated phishing emails, respectively; ATN_i = attention to phishing cues; TP_i = task performance; $CONTROL_{ki}$ = 10 control variables.
- β_k = regression coefficients; in generalized estimating equations, the regression coefficients represent the population-average change in the dependent variable for a one-unit change in the independent variable.

Table 7 shows the results, respectively, of the GEE models for task-related and task-unrelated phishing emails. Model 2 fit the data much better than Model 1, according to Quasi-likelihood under Independence Model Criterion (QIC) and Corrected Quasi-likelihood Under Independence Model Criterion (QICC).

We can see that for both types of phishing emails, attention to phishing cues positively affected antiphishing performance (0.218, Wald $\chi 2$ (1) = 5.206, $p < 0.05$, for task-related phishing; 0.309, Wald $\chi 2$ (1) = 8.839, $p < 0.01$, for task-unrelated phishing). This result suggested that attention to phishing cues positively affected antiphishing performance for a variety of phishing emails. Meanwhile, for task-related phishing emails, task performance negatively affected antiphishing performance (-0.417, Wald $\chi 2$ (1) = 17.134, $p < 0.001$). Attention to phishing cues moderated the relationship between task performance and antiphishing performance but was marginally significant (0.152, Wald $\chi 2$ (1) = 2.787, $p < 0.10$). However, for task-unrelated phishing emails, task performance did not significantly affect antiphishing performance (-0.153, Wald $\chi 2$ (1) = 2.199, $p = 0.138$). Attention to phishing cues did not significantly moderate the relationship between task and antiphishing performance (0.018, Wald $\chi 2$ (1) = 0.038, $p = 0.845$). The results based on GEE suggested that task performance negatively affected antiphishing performance for task-related phishing emails but did not necessarily affect antiphishing performance for those unrelated to tasks. Moreover, attention to phishing cues moderated the negative relationship between task performance and antiphishing performance for task-related phishing emails. Attention to phishing cues didn't moderate the relationship between task performance and antiphishing performance for task-unrelated phishing emails. Based on the previous results, the second part of H7 was supported, and H8 was marginally supported.

Table 7. Results of Generalized Estimating Equations

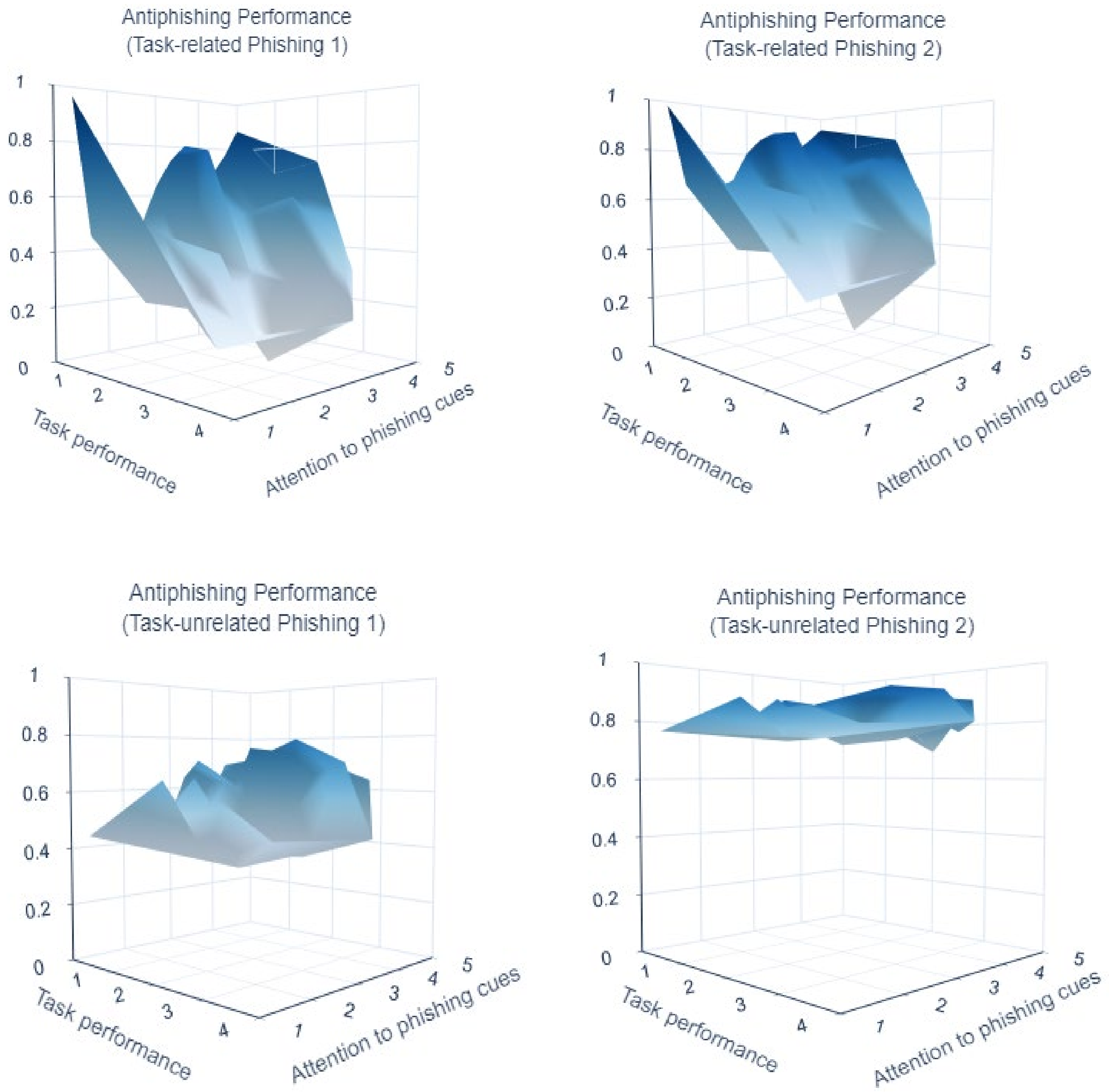
	Task-related Phishing		Task-unrelated Phishing	
	Model 1	Model 2	Model 1	Model 2
Intercept	0.317^{**}	0.176	0.605^{***}	0.654^{***}
Level 1				
TIME 2	0.770^{***}	0.865^{***}	1.371^{***}	1.452^{***}
TIME 1	0.000^{a}	0.000^{a}	0.000^{a}	0.000^{a}
Level 2				
AGE		-0.079		0.311^{*}
GEN		$0.347^{\dagger}$		0.014
DSK		0.039		0.000
LAP		-0.150		-0.070
TAB		0.039		-0.091
PHN		-0.267^{**}		-0.048
LOAD		0.306^{**}		-0.021
EXP		-0.223^{**}		-0.223^{*}
THR		0.113		$-0.204^{\dagger}$
ANX		-0.279^{**}		0.257^{*}
ATN		0.218^{*}		0.309^{**}
TP		-0.417^{***}		-0.153
ATN*TP		$0.152^{\dagger}$		0.018
Model Fit				
QIC	843.464	799.051	691.195	681.147
QICC	843.464	795.620	691.195	681.004

Notes:

- N (Level 1) = 674 and N (Level 2) = 337
- a: Set to 0
- GEN = gender; DSK = desktop use; LAP = laptop use; TAB = tablet use; PHN = smartphone use; LOAD = email load; EXP = prior phishing experience; THR = perceived threat; ANX = phishing anxiety; ATN = attention to phishing cues; AP = antiphishing performance; TP = task performance.
- $^{\dagger}p < 0.10$, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$ (two-tailed).

We also plotted the relationships between task performance, attention to phishing cues, and antiphishing performance for each of the task-related and task-unrelated phishing emails in Figure 6. The figure shows that for both task-related phishing emails, antiphishing performance is better when task performance is low or when attention to phishing cues is high. However, for task-unrelated phishing

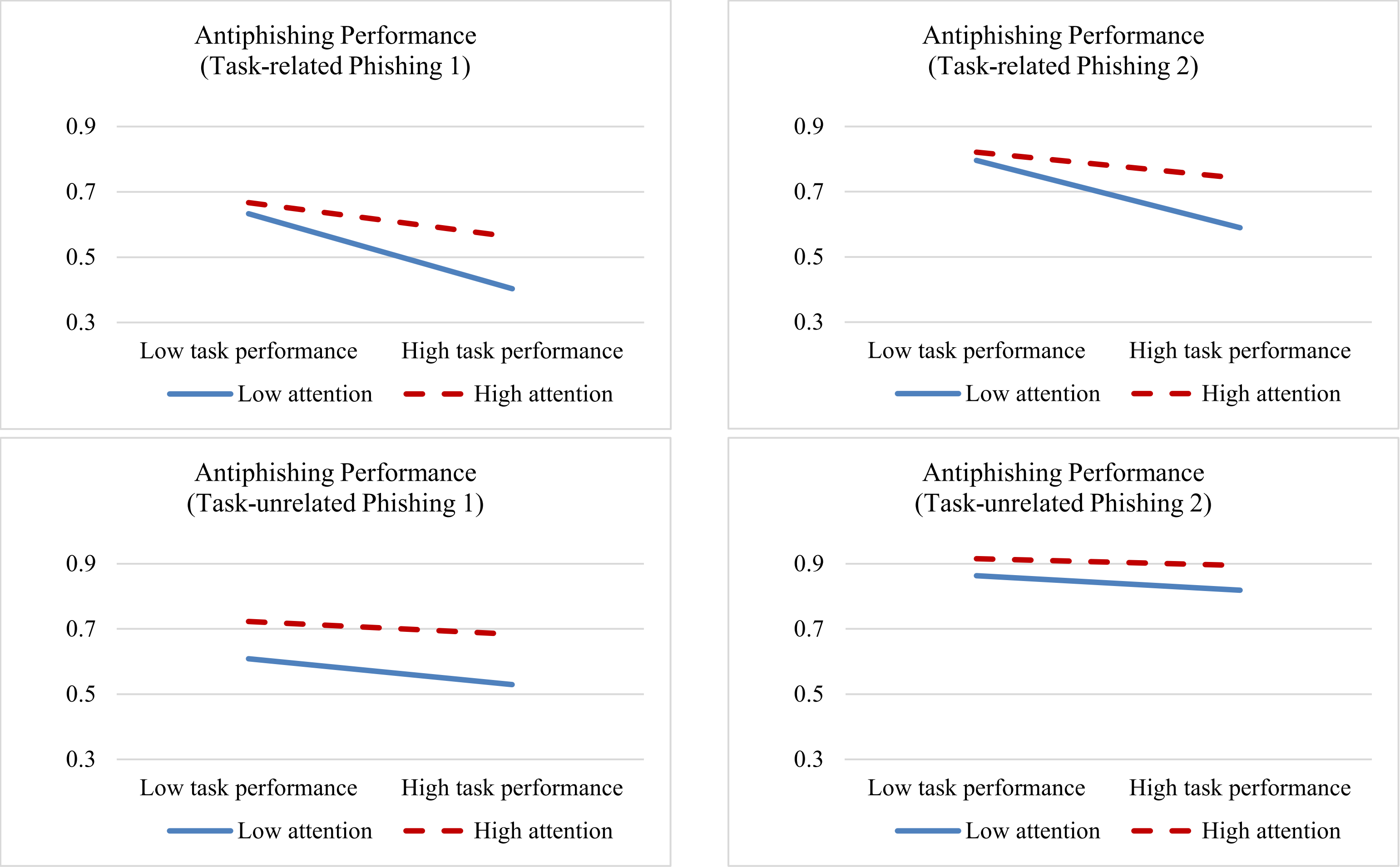
Figure 6. 3-D Plots of GEE Results



Notes:

- Antiphishing performance is the predicted probability of avoiding a phishing email.
- Task-related phishing 1 and 2 are phishing 1 and 4 in Appendix B. Task-unrelated phishing 1 and 2 are phishing 2 and 3 in Appendix B.
- The figures are based on GEE models with research variables in their original scales.

Figure 7. Plots of the Interaction Effect



Notes:

- Antiphishing performance is the average predicted probability of avoiding phishing emails in the high/low task performance and attention groups.
- If a person is in a high (low) task performance/attention group, that person’s score on task performance/attention is above (below) the mean score.

emails, antiphishing performance does not decline even if task performance increases. Instead, antiphishing performance moderately increases as attention increases.

To better illustrate the effect of task performance, attention to phishing cues, and the interaction on antiphishing performance, we also plotted the two-dimensional graphs as shown in Figure 7. Figure 7 shows that for the two task-related phishing emails, participants reporting high attention demonstrate better antiphishing performance than participants with low attention. Moreover, participants' antiphishing performance decreases significantly as their task performance increases. Antiphishing performance drops more severely as task performance increases for participants with low attention than for participants with high attention. For the two task-unrelated phishing emails, participants with high attention still show better antiphishing performance. However, antiphishing performance does not decrease significantly as task performance increases. Attention to phishing cues has no effect on the change in antiphishing performance as task performance increases. In summary, Figure 7 indicates that attention to phishing cues is positively associated with antiphishing performance for both task-related and task-unrelated phishing emails. However, task performance is negatively associated with antiphishing performance only for task-related phishing emails, but attention to phishing cues moderates this negative relationship.

6. Discussion and Conclusion

The objective of this study is to develop and test a model of antiphishing, especially in its relation to task performance in checking legitimate work emails. Drawing on the model of safety performance and a goal-oriented framework, we developed a model that highlights the important roles of antiphishing self-efficacy, antiphishing climate, antiphishing motivation, attention to phishing cues, and task performance in explaining antiphishing behavior. We tested the model by using a field experiment with a survey questionnaire in which four legitimate emails and four phishing emails were sent to 357 participants. The data were analyzed through SEM and GEE. As hypothesized, antiphishing climate, antiphishing self-efficacy, and antiphishing motivation played important roles in regulating attention to phishing cues. In addition, the results showed that for task-related phishing emails, task performance in keeping up with work emails was negatively related with antiphishing performance. Attention to phishing cues was positively associated with antiphishing performance; and for task-related phishing emails, it also decreased the negative relationship between task performance and antiphishing performance. These findings shed light on the potential of a goal-oriented framework in studying antiphishing behavior and have practical implications about the conflict between antiphishing performance and task performance.

6.1. Theoretical Contributions

The major contribution of this study is to combine the goal-oriented framework and the model of safety performance for a logical and powerful theoretical account of antiphishing in a multigoal context. This study reveals the similarity of the safety model with a goal-oriented framework and further extends the safety model to better explain antiphishing, especially when the goal of detecting phishing emails conflicts with other daily work-related goals such as keeping up with work emails. Despite the existence of multiple goals in everyday life, previous researchers have examined antiphishing with little consideration for people's pursuit of daily work goals. The present study is meaningful in its taking an initial step toward a better understanding of antiphishing vis-à-vis task operation in individual's daily work. In summary, this study demonstrates the potential of a goal-oriented approach to antiphishing research, and it is expected to further enrich our insights into the complex process of goal setting, goal pursuit, and goal achievement in a conflicting environment of antiphishing and daily work activities.

Drawing on the goal-oriented framework, we have demonstrated the potential conflicts that can arise between achieving the goal of combating phishing emails and achieving other daily work goals. Consistent with the proposed model, our findings showed an increase in task performance in daily work actually led to a decrease in antiphishing performance. Specifically, we found that the negative relationship between task performance and antiphishing performance was significant only for task-related phishing messages and not for phishing messages unrelated to the daily task. These findings further bolster our goal-oriented view that some (e.g., task-related) antiphishing activities conflict more with a certain type of daily work than other (e.g., task-unrelated) antiphishing activities.

Consistent with the goal-oriented framework, attention to phishing cues, or goal-directed efforts regarding combating phishing emails, was shown to improve antiphishing performance. Moreover, the present study is the first to theoretically propose and empirically show that attention serves as a moderator between task performance and antiphishing performance. Drawing on the notion of attentional resources, we theorize that people can be deliberately attentive, if necessary, to achieve simultaneously the goal of combating phishing emails and the goals of daily work. As a result, the negative effect of task performance on antiphishing performance would decrease with the increase in attention to phishing cues. Our findings show that attention to phishing cues indeed mitigates the negative relationship between task performance and antiphishing performance in the case of task-related phishing messages. We also found that attention fully mediated the impacts of its antecedents on antiphishing performance. Overall, this study contributes to IS research by highlighting the role of attention to phishing cues in mediating the impact of motivation, self-efficacy, and climate on antiphishing performance and especially its moderating effect on the conflicting relationship between task operations in daily work and antiphishing.

Few studies have attempted to provide a theoretical account of the role of perceived climate in regulating individuals' antiphishing behavior. However, much research has been focused on individual-level self-efficacy, emotions, and judgment. Drawing on the model of safety performance, we developed a theoretical framework that describes how the antiphishing climate affects individuals' reactions to phishing attacks. Because of the dynamically evolving nature of phishing attacks that differ from traditional brick-and-mortar environments, we theorized that an antiphishing climate would have negative effects—instead of the positive effects implied by the safety model—on antiphishing motivation and antiphishing behavior. Although in this study the climate factor was not a determinant of motivation, it was shown to have a negative impact on antiphishing behavior (i.e., attention to phishing cues). These results provide empirical support for our claim that people tend to be inattentive to phishing cues when they feel that their organization is well prepared for potential phishing attacks. This study contributes to antiphishing research by showing theoretical and empirical evidence for the counterintuitive role of perceived climate in determining individuals' reactions to phishing emails.

Taken together, this study is an initial attempt to uncover the importance of goal conflicts in the study of antiphishing and to caution researchers against examining antiphishing without taking into account broader contextual considerations.

6.2. Practical Implications

This study demonstrates an interesting finding that an antiphishing climate is negatively related with attention to phishing cues. That is, those who consider their climate cooperative tend to place too much confidence in their organization, and ironically, they neglect their own responsibilities and due diligence. A managerial insight that can be acquired from these findings is that organizations should not overstate or hide their actual antiphishing capabilities. In particular, we suggest practitioners make transparent to their members the current situation related to failed antiphishing efforts and their negative consequences. For example, organizations may issue periodic reports on the number of phishing emails that successfully penetrate their networks as well as the number of actual victims. Such organizational actions will make members more aware of the actual risks associated with phishing-related threats and feel personally responsible in properly responding to future attacks. As a whole, organizations should help members have realistic expectations about what organizations can do for safe computing and understand that some phishing attacks cannot be avoided despite strong technical and educational countermeasures.

This study also demonstrates that those who click more on legitimate work-relevant emails tend to click more on task-related phishing emails. These findings paradoxically indicate that hardworking people who are eager to go extra miles for better outcomes are likelier to fall victim to phishing attacks. Consequently, we recommend managers offer organizational assistance to these devoted workers. For example, firms may offer mindfulness training programs as recommended by Jensen et al. (2017) and implement technical solutions such as two-factor authentication within the email systems. In general, this study shows that the most committed workers can be the weakest link in information security. Thus, managers should provide them with managerial and technical support to help them better discern between legitimate and phishing emails.

6.3. Limitations and Future Research

Some limitations of this paper need to be mentioned. First, in examining the dynamics between antiphishing and task operation, we did not measure task climate, motivation, or task behavior while focusing solely on task performance. A more complete picture of antiphishing in daily work can be drawn when the effects of task-related variables not included in this study are fully understood. Second, although we strived to control several factors that are believed important in antiphishing in daily work (e.g., anxiety, perceived threats, and IT use), we may have failed to include in our model still other variables that are potentially important. Thus, caution should be exercised in interpreting our findings until they are validated in a more comprehensive model. Third, given that our data were focused on phishing and legitimate messages in relation to course work at college, it remains to be seen whether our findings can be generalized to an actual business setting. Further research needs to explore employees who work in an organization, and it will enrich our understanding about conflicts between individuals' security-related goal and their daily work goals. Fourth, although we used eight phishing and legitimate messages to accurately evaluate antiphishing and task performance, our results might be different if we used other messages. Fifth, although we found out a counterintuitive, negative impact of antiphishing climate, we need to further examine the boundary conditions of this finding. For example, the effectiveness of the organization's practices should be included in future studies in order to fully understand the impact of antiphishing climate. In addition, although we tried to eliminate common method variance by measuring our dependent and independent variables separately, it would be better to further validate our findings by using a more sophisticated design (e.g., to also measure antiphishing climate and attention to phishing cues at different times). Our findings should therefore be interpreted with care until the above issues have been fully studied. Lastly, in the current study, we recorded individuals' clicking behavior on fraudulent links in phishing emails to measure their antiphishing performance. Future research is encouraged to examine different stages of interacting with phishing emails (e.g., opening the phishing emails, clicking on the links and entering personal information into the fake websites). A detailed examination of the antiphishing process can provide more insights into how the goal of combating phishing emails conflicts with other work-related goals.

This study opens up several exciting avenues for further research. First of all, although the current study examines how task performance in keeping up with work emails influences antiphishing performance, further research is needed to explore whether and how other tasks can influence antiphishing performance. Moreover, future research is needed to study how task-related perceptions and behavior are interrelated with antiphishing-related perceptions and behavior. In particular, we found task performance negatively affected antiphishing performance, but it remains to be seen whether task-related perceptions and behavior would similarly influence antiphishing-related perceptions and behavior. To uncover the interrelated nature of work-related tasks and antiphishing, researchers are encouraged to study carefully their interrelationships not only at the performance level but also at the perception and behavioral levels.

We also suggest researchers carefully examine how goals are set as a way for them to gain a better understanding of goal pursuit and goal achievement in the context of antiphishing. Research indicates that goal achievement varies with whether such goals are set by an actor or enforced by an organization (Loock et al., 2013). Whereas the current study assumes that a desired level of antiphishing is loosely set by an individual, an organization can alternatively punish the lack of a proper defense against phishing attacks. In such a mandatory situation, antiphishing motivation, behavior, and performance are unlikely to be identical to those we examined in this study. More interestingly, an organization can also demand a specific level of task performance. The organizational requirements of antiphishing and task performance are likely to substantially transform the way that people work. Further research is thus required to investigate how mandatory and voluntary goals can differently regulate antiphishing behavior.

6.4. Conclusions

Users of information technology need to be alert to potential phishing attacks while responding efficiently to their work-related messages. Drawing on the safety model and a goal-oriented framework, we developed a theoretical framework that explains how IT users deal with various phishing and legitimate emails. We found that attention to phishing cues is a key to avoiding not only task-related phishing attacks but also task-unrelated phishing attacks. More interestingly, this study shows that attention to phishing cues can moderate the harmful effect of task performance on antiphishing performance in the case of task-related phishing emails. Our model is a powerful tool for offering a theoretical account of antiphishing, especially when it is at odds with task operations in daily work. Researchers are encouraged to examine the issues of antiphishing within a more realistic and everyday work environment. We hope that the conceptual model proposed in this study will be useful for such endeavors.

References

Abbasi, A., Dobolyi, D., Vance, A., & Zahedi, F. M. (2021). The phishing funnel model: A design artifact to predict user susceptibility to phishing websites. *Information Systems Research, 32*(2), 410-436.

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

Ajzen, I. (1985). From intentions to actions: A theory of planned behavior. In J. Kuhl, & J. Beckmann (Eds.), *Action control: From cognition to behavior*. New York: Springer Verlag, 11-39.

Arachchilage, N. A. G., & Love, S. (2013). A game design framework for avoiding phishing attacks. *Computers in Human Behavior, 29*(3), 706-714.

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

Arachchilage, N. A. G., Love, S., & Beznosov, K. (2016). Phishing threat avoidance behaviour: An empirical investigation. *Computers in Human Behavior, 60*, 185-197.

Bagozzi, R. P. (2007). The legacy of the technology acceptance model and a proposal for a paradigm shift. *Journal of the Association for Information Systems, 8*(4), 244-254.

Bagozzi, R. P., & Dholakia, U. (1999). Goal setting and goal striving in consumer behavior. *Journal of Marketing, 63(special issue)*, 19-32.

Bagozzi, R. P., & Yi, Y. (1988). On the evaluation of structural equation models. *Journal of the Academy of Marketing Science, 16*(1), 74-94.

Bagozzi, R. P., Dholakia, U. M., & Basuroy, S. (2003). How effortful decisions get enacted: The motivating role of decision processes, desires, and anticipated emotions. *Journal of Behavioral Decision Making, 16*(4), 273-295.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84*(2), 191-215.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: W. H. Freeman.

Barling, J., Loughlin, C., & Kelloway, E. K. (2002). Development and test of a model linking safety-specific transformational leadership and occupational safety. *Journal of Applied Psychology, 87*(3), 488-496.

Bissell, K., LaSalle, R., & Cin, P. D. (2019). Ninth annual cost of cybercrime study. Dublin: Ponemon Institute.

Brown, K. W., & Ryan, R. M. (2003). The benefits of being present: Mindfulness and its role in psychological well-being. *Journal of Personality and Social Psychology*, *84*(4), 822-848.

Brown, K. W., Ryan, R. M., & Creswell, J. D. (2007). Mindfulness: Theoretical foundations and evidence for its salutary effects. *Psychological Inquiry, 18*(4), 211-237.

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

Chin, W. W., Gopal, A., & Salisbury, W. D. (1997). Advancing the theory of adaptive structuration: The development of a scale to measure faithfulness of appropriation. *Information Systems Research, 8*(4), 342-367.

Chowdhury, N. H., Adam, M. T., & Skinner, G. (2018). The impact of time pressure on human cybersecurity behavior: An integrative framework. *2018 26th International Conference on Systems Engineering (ICSEng)*, New York, NY: Institute of Electrical and Electronics Engineers, 1-10.

Clarke, S. (2006). The relationship between safety climate and safety performance: A meta-analytic review. *Journal of Occupational Health Psychology, 11*(4), 315–327.

Dincelli, E., & Chengalur-Smith, I. (2020). Choose your own training adventure: Designing a gamified SETA artefact for improving information security and privacy through interactive storytelling. *European Journal of Information Systems, 29*(6), 669-687.

Fishbach, A., & Choi, J. (2012). When thinking about goals undermines goal pursuit. *Organizational Behavior and Human Decision Processes, 118*(2), 99-107.

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research, 18*(1), 39-50.

Gefen, D. (2002). Reflections on the dimensions of trust and trustworthiness among online consumers. *ACM SIGMIS Database: the DATABASE for Advances in Information Systems, 33*(3), 38-53.

Gist, M. E., & Mitchell, T. R. (1992). Self-efficacy: A theoretical analysis of its determinants and malleability. *Academy of Management Review, 17*(2), 183-211.

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

Goes, P. B., Guo, C., & Lin, M. (2016). Do incentive hierarchies induce user effort? Evidence from an online knowledge exchange. *Information Systems Research, 27*(3), 497-516.

Gollwitzer, P .M. (1996). Benefits of planning. In P. M. Gollwitzer, & J. A. Bargh (Eds.), *The psychology of action: Linking cognition and motivation to behavior*. New York, NY: Guilford Press, 287-312.

Heck, R. H., Thomas, S., & Tabata, L. (2012). *Multilevel modeling of categorical outcomes using IBM SPSS*. New York, NY: Routledge Academic.

Hong, J. (2012). The state of phishing attacks. *Communications of the ACM, 55*(1), 74-81.

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), 597-626.

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

Kankanhalli, A., Tan, B. C., & Wei, K. K. (2005). Contributing knowledge to electronic knowledge repositories: An empirical investigation. *MIS Quarterly, 29*(1), 113-143.

Kim, S. S. (2009). The integrative framework of technology use: An extension and test. *MIS Quarterly, 33*(3), 513-537.

Kumaraguru, P., Cranshaw, J., Acquisti, A., Cranor, L., Hong, J., Blair, M. A., & Pham, T. (2009). School of phish: A real-world evaluation of anti-phishing training. *Proceedings of the 5th Symposium on Usable Privacy and Security*, New York, NY: Association for Computing Machinery, 1-12.

Kumaraguru, P., Rhee, Y., Acquisti, A., Cranor, L. F., Hong, J., & Nunge, E. (2007a). Protecting people from phishing: The design and evaluation of an embedded training email system. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, New York, NY: Association for Computing Machinery, 905-914.

Kumaraguru, P., Rhee, Y., Sheng, S., Hasan, S., Acquisti, A., Cranor, L. F., & Hong, J. (2007b). Getting users to pay attention to anti-phishing education: Evaluation of retention and transfer. *Proceedings of the Anti-phishing Working Groups 2nd Annual eCrime Researchers Summi*t, New York, NY: Association for Computing Machinery, 70-81.

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

Loock, C. M., Staake, T., & Thiesse, F. (2013). Motivating energy-efficient behavior with green IS: An investigation of goal setting and the role of defaults. *MIS Quarterly, 37*(4), 1313-1332.

Louro, M. J., Pieters, R., & Zeelenberg, M. (2007). Dynamics of multiple goal pursuit. *Journal of Personality and Social Psychology, 93*(2), 174-193.

MacKillop, J., & Anderson, E. J. (2007). Further psychometric validation of the mindful attention awareness scale (MAAS). *Journal of Psychopathology and Behavioral Assessment, 29*(4), 289-293.

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

Mayer, P., Gerber, N., McDermott, R., Volkamer, M., & Vogt, J. (2017). Productivity vs security: Mitigating conflicting goals in organizations. *Information and Computer Security, 25*(2), 137-151.

Moody, G. D., Siponen, M., & Pahnila, S. (2018). Toward a unified model of information security policy compliance. *MIS Quarterly, 42*(1), 285-311.

Morrow, S. L., McGonagle, A. K., Dove-Steinkamp, M. L., Walker Jr, C. T., Marmet, M., & Barnes-Farrell, J. L. (2010). Relationships between psychological safety climate facets and safety behavior in the rail industry: A dominance analysis. *Accident Analysis and Prevention, 42*(5), 1460-1467.

Nahrgang, J. D., Morgeson, F. P., & Hofmann, D. A. (2011). Safety at work: A meta-analytic investigation of the link between job demands, job resources, burnout, engagement, and safety outcomes. *Journal of Applied Psychology, 96*(1), 71–94.

Neal, A., & Griffin, M. A. (2006). A Study of the lagged relationships among safety climate, safety motivation, safety behavior, and accidents at the individual and group levels. *Journal of Applied Psychology, 91*(4), 946–953.

Neal, A., Griffin, M. A, & Hart, P. M. (2000). The impact of organizational climate on safety climate and individual behavior. *Safety Science, 34*, 99-109.

Nguyen, C., Jensen, M. L., Durcikova, A., & Wright, R. T. (2021). A comparison of features in a crowdsourced phishing warning system. *Information Systems Journal, 31*(3), 473-513.

Nicolaou, A. I., & McKnight, D. H. (2006). Perceived information quality in data exchanges: Effects on risk, trust, and intention to use. *Information Systems Research, 17*(4), 332-351.

Parker, C. P. (2003). Relationships between psychological climate perceptions and work outcomes: A meta-analytic review. *Journal of Organizational Behavior, 24*(4), 389-416.

Pavlou, P. A. (2003). Consumer acceptance of electronic commerce: Integrating trust and risk with the technology acceptance model. *International Journal of Electronic Commerce, 7*(3), 101-134.

Polites, G. L., & Karahanna, E. (2012). Shackled to the status quo: The inhibiting effects of incumbent system habit, switching costs, and inertia on new system acceptance. *MIS quarterly, 36*(1), 21-42.

Ray, S., Kim, S. S., & Morris, J. G. (2014). The central role of engagement in online communities. *Information Systems Research, 25*(3), 528-546.

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist, 55*(1), 68-78.

Schmitz, K., Teng, J. T., & Webb, K. (2016). Capturing the complexity of malleable IT use: Adaptive structuration theory for individuals. *MIS Quarterly, 40*(3), 663-686.

Segars, A. H. (1997). Assessing the unidimensionality of measurement: A paradigm and illustration within the context of information systems research. *Omega, 25*(1), 107-121.

Sekerka, L. E., & Bagozzi, R. P. (2007). Moral courage in the workplace: Moving to and from the desire and decision to act. *Business Ethics: A European Review, 16*(2), 132-149.

Sharma, R., Yetton, P., & Crawford, J. (2009). Estimating the effect of common method variance: The method—method pair technique with an illustration from TAM Research. *MIS Quarterly, 33*(3), 473-490.

Sheldon, K. M., & Elliot, A. (1999). Goal striving, need satisfaction and longitudinal well-being, the self-concordance model. *Journal of Personality and Social Psychology, 76*(3), 486-497.

Sheng, S., Holbrook, M., Kumaraguru, P., Cranor, L. F., & Downs, J. (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*, New York, NY: Association for Computing Machinery, 373-382.

Sheng, S., Magnien, B., Kumaraguru, P., Acquisti, A., Cranor, L. F., Hong, J., & Nunge, E. (2007). Anti-phishing phil: The design and evaluation of a game that teaches people not to fall for phish. *Proceedings of the 3rd Symposium on Usable Privacy and Security*, New York, NY: Association for Computing Machinery, 88-99.

Singer, S., Lin, S., Falwell, A., Gaba, D., & Baker, L. (2009). Relationship of safety climate and safety performance in hospitals. *Health Services Research, 44*(2), 399-421.

Stajkovic, A. D., & Luthans, F. (1998). Self-efficacy and work-related performance: A meta-analysis. *Psychological Bulletin, 124*(2), 240-261.

Strayer, D. L., & Johnston, W. A. (2001). Driven to distraction: Dual-task studies of simulated driving and conversing on a cellular telephone. *Psychological Science, 12*(6), 462-466.

Strayer, D. L., Drews, F., & Johnston, W. A. (2003). Cell phone-induced failures of visual attention during simulated driving. *Journal of Experimental Psychology: Applied, 9*(1), 23-32.

Thompson, N., McGill, T. J., & Wang, X. (2017). "Security begins at home": Determinants of home computer and mobile device security behavior. *Computers and Security, 70*, 376-391.

Vancouver, J. B., Thompson, C. M., & Williams, A. A. (2001). The changing signs in the relationships among self-efficacy, personal goals, and performance. *Journal of Applied Psychology, 86*(4), 605-620.

Vancouver, J. B., Thompson, C. M., Tischner, E. C., & Putka, D. J. (2002). Two studies examining the negative effect of self-efficacy on performance. *Journal of Applied Psychology, 87*(3), 506-516.

Vancouver, J. B., Weinhardt, J. M., & Schmidt, A. M. (2010). A formal, computational theory of multiple-goal pursuit: Integrating goal-choice and goal-striving processes. *Journal of Applied Psychology, 95*(6), 985–1008.

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly, 27*(3), 425-478.

Vinodkumar, M. N., & Bhasi, M. (2009). Safety climate factors and its relationship with accidents and personal attributes in the chemical industry. *Safety Science, 47*(5), 659-667.

Vishwanath, A. (2015). Habitual facebook use and its impact on getting deceived on social media. *Journal of Computer-Mediated Communication, 20*, 83-98.

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), 576-586.

Wang, J., Li, Y., & Rao, H. R. (2017). Coping responses in phishing detection: An investigation. *Information Systems Research, 28*(2), 378-396.

Watson, J. M., & Strayer, D. L. (2010). Supertaskers: Profiles in extraordinary multitasking ability. *Psychonomic Bulletin and Review, 17*(4), 479-485.

Wright, R. T. (2014). Influence techniques in phishing attacks:An examination of vulnerability and resistance. *Information Systems Research, 25*(2), 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), 273-303.

Yeo, G. B., & Neal, A. (2006). An examination of the dynamic relationship between self-efficacy and performance across levels of analysis and levels of specificity. *Journal of Applied Psychology, 91*(5), 1088-1101.

Appendix A: Prior Studies on Antiphishing Performance

Table 8. Prior Studies on Antiphishing Performance

Study	Method	Sample	Focus of the Study	Theory	Research Factors
Kumaraguru et al. 2007a	Lab experiment	30 participants within a university and its neighborhood	Evaluating the effectiveness of an embedded email training system for improving phishing detection performance	Learning science theory	Phishing detection performance, interventions, gender, email load, age
Kumaraguru et al. 2007b	Lab experiment	42 participants in and near a university	The retention and transfer of the effectiveness of an embedded email training system for improving phishing detection performance	Learning science theory	Phishing detection performance, interventions, gender, email load, age, cognitive reflection
Sheng et al. 2007	Lab experiment with a survey	42 participants recruited around a university	Testing the effectiveness of an antiphishing educational computer game to help people detect fake websites	Learning science theory, signal detection theory	Detection performance of both phishing and legitimate websites, user confidence, user satisfaction, demographic variables
Sheng et al. 2010	Survey with a role-playing task	1,001 online participants	Testing the effectiveness of educational cartoons, computer games and web-based training materials	N/A	Detection performance of phishing and legitimate emails, previous online experience, age, gender, education, technical knowledge, and risk aversion
Wright and Marett 2010	Field experiment with survey	299 undergraduate students	Studying individual experiential and dispositional factors that lead to susceptibility to phishing	Interpersonal deception theory	Deception success, computer self-efficacy, security knowledge, web experience, trust, perceived risk, and suspicion of humanity
Vishwanath et al. 2011	Survey after real phishing attacks	321 undergraduate students	Testing factors that lead to susceptibility to phishing attacks	O-S-I-R model	Attention to email content and layout, elaboration, email involvement, email load, domain specific knowledge, computer self-efficacy

Arachchilage and Love 2013	Survey	151 undergraduate students	Factors that should be addressed for an antiphishing game design	Technology threat avoidance theory	Avoidance behavior, avoidance motivation, perceived threat, safeguard cost, safeguard effectiveness, self-efficacy, perceived severity, perceived susceptibility
Arachchilage and Love 2014	Survey	161 undergraduate students	Antecedents of self-efficacy related to gain antiphishing knowledge	Technology threat avoidance theory	Procedural knowledge, conceptual knowledge, self-efficacy, avoidance motivation, avoidance behavior, demographics
Wright et al. 2014	Field experiment	2,624 university students	Studying the effectiveness of different influence techniques to deceive people	Persuasion and motivation theory	Liking, social proof, consistency, authority, scarcity
Vishwanath 2015	Field experiment with survey	155 undergraduate students	Studying the influence of habitual Facebook use on antiphishing behavior on Facebook	Habit and addiction	Frequency of Facebook use, Facebook habit strength, deficient self-regulation, number of Facebook friends, concern for privacy, attitudinal commitment, phishing avoidance behavior
Arachchilage and Love 2016	Think-aloud study with a pre- and post- test	20 computer science undergraduate students	Testing the usability and effectiveness of an antiphishing educational mobile game	Technology threat avoidance theory	Detection performance of illegal websites, user satisfaction about the game
Goel et al. 2017	Field experiment	7,225 undergraduate students	Testing the effectiveness of different deceptive strategies	Heuristic-systematic processing model	Gains or loss, contextualization, general motives
Jensen et al. 2017	Field experiment	355 participants within a university	Testing the effectiveness of mindfulness-based training in comparison with rule-based training in detecting phishing emails	Mindfulness theory	Phishing detection performance, email mindfulness, trust, perceived risk, computer self-efficacy, phishing detection expertise

Wang et al. 2017	Survey experiment	547 consumers within USA	Testing factors that influence phishing detection accuracy and effort	Extended parallel process model	Phishing detection accuracy, detection effort, coping adaptiveness, phishing anxiety, perceived detection efficacy, perceived susceptibility, perceived severity, age, gender, education, prior victimization, income, internet experience, email load, number of credit cards, dispositional optimism, familiarity with the sender
Dincelli & Smith 2020	Longitudinal randomized controlled experiment	1,718 employees on MTurk	Testing the effectiveness of a gamified security awareness, education and training artefact	Gamification	Game design elements, online self-disclosure behavior, attitudes towards self-disclosure, intentions towards self-disclosure, memorability of the education, use experience of the education
Abbasi et al. 2021	Longitudinal field experiment	1,278 employees in organizations	Testing the effectiveness of a warning system based on a phishing funnel model	The funnel concept	Tool information, tool perception, threat characteristics, threat perceptions, demographics, prior web experiences, phishing susceptibility
Nguyen et al. 2021	Survey experiment	438 students	Testing the impact of different features of a crowdsourced warning system on individual phishing avoidance behavior	Crowdsourcing	Number of reports, report source, accuracy rate, disclosure of accuracy rate, adherence to warning recommendations, anxiety with the warning system, phishing avoidance behavior

Appendix B: Legitimate and Phishing Emails

[Legitimate Email 1]
From: TA of the Course <xxx@xxx.edu>
Date: XXX
To:
Subject: Course name: Come to Know More about Your Instructor

Hi All,

If you want to check the bio of the instructor, please visit this Link. This information cannot be found on the course website.

Thanks,

XXX

TA for the course

[Phishing 1]
From: Grayson Wright <administrator@remoteemail.net>
Date: XXX
To:
Subject: Course message from XXX course

Dear students,

There is 1 course message for you from XXX course.
Click to view

Grayson Wright

[Legitimate Email 2]
From: TA of the Course <xxx@xxx.edu>
Date: XXX
To:
Subject: Course name: Getting to Know Your TA

Hi All,

If you are interested in your TA's profile, please view it through the Link. You can't find the information from the course website.

Thanks!

[Phishing 2]
From: The University's School of Business <education@badgeremail.org>
Date: XXX
To:
Subject: Enjoy a Gift Card from the University's School of Business

Dear student's name,

Thank you for participating in the XXX event! Please click here to redeem a $50.00 University Bookstore Gift Card. If you are not able to redeem the card, please contact Jane Brennan at JaneBrennan@xxx.edu for immediate assistance.

The University's School of Business

The Logo of The University's
School of Business

[Legitimate Email 3]
From: TA of the Course <xxx@xxx.edu>
Date: XXX
To:
Subject: Danger of Business Analytics

Hi All,

Now that you are done with the first exam and feel more confident about Business Analytics, we would like to caution you against its overuse in reality. Let's see a classic example of BAD Business Analytics through the Link. It is a silly cartoon but seems appropriate especially after the serious exam :-)

Have a nice weekend!

[Phishing 3]
From: IT Support <admin@it-email-services.com>
Date: XXX
To:
Subject: ID Security Alert

Student's name,
Someone else was trying to use your XX ID to sign into Google via a web browser.

Date and Time: XXX

Browser: Chrome

Location: XXX

If you believe someone may be trying to access your account, please click <Here>.

Sincerely,
Technical Support Team

[Legitimate Email 4]
From: TA of the Course <xxx@xxx.edu>
Date: XXX
To:
Subject: SQL for DB lecture

Dear students,

Please find a PDF file on SQL syntax through this Link. It will be discussed in the DB lecture this week. You will not be able to find this PDF file on the course website.

Best

[Phishing 4]
From: The University's Learning Management System <no-reply@link-it.us>
Date: XXX
To:
Subject: Learning Management System Announcement

You have an announcement from your course instructor. Please click here to see the message on the Learning Management System.

The Logo of
the Learning Management System

Essay 2: Roles of Feedback and Phishing Characteristics in Identifying Phishing Scams: Conceptual Model and Three Experiments

Abstract

Because phishing attacks often exploit individuals' inexperience in detecting them, it is important for managers to provide workers with proper feedback on their reactions to phishing scams. However, little is known about what types of feedback are more effective in facilitating antiphishing behavior and performance. The objective of this study is to develop and test a model on the effect of feedback on decision avoidance and detection accuracy in antiphishing training. Our model provides a theoretical account of how the relationships between phishing characteristics (e.g., phishing cue saliency) and antiphishing outcomes change with feedback characteristics (e.g., feedback quantity). Drawing on theories of feasibility, we propose that perceived detection feasibility is a key factor that intervenes between antecedents (feedback and phishing characteristics) and outcomes (decision avoidance and detection accuracy). The proposed model further extends the notion of feasibility with the concept of skill acquisition to explain how antecedents interact to influence outcomes. To empirically test the model, we performed three experiments with 514 subjects in the United States. Our results indicate that feedback with technical information is better than feedback with abstract information for correctly discerning between phishing and legitimate emails. However, feedback with abstract information tends to make individuals more cautious to make decision about an email's legitimacy. We also show that perceived detection feasibility is essential for a better understanding of antiphishing behavior and performance. Most important, we show interesting interaction effect between feedback quantity and phishing cue saliency on antiphishing performance.

Keywords: phishing, antiphishing training, feedback, phishing cue saliency, perceived detection feasibility, skill acquisition, decision avoidance, detection accuracy, experiments, hierarchical linear modeling

1. Introduction

Phishing attacks can cause an enormous amount of financial and data loss in organizations. Industry estimates indicate that phishing is the number one cause of data breaches (Verizon, 2020). Recent studies show that phishing attacks targeted nearly 90% of organizations in 2019, and these phishing attacks resulted in losses of more than $1.8 billion (FBI, 2019; Proofpoint, 2020). Because phishing attacks often exploit individuals' inexperience with phishing scams, it is important for managers to train employees to be aware of phishing and how to detect it (Al-Daeef et al., 2017; Arachchilage et al., 2016; Wright et al., 2014). Organizations often offer phishing awareness programs that include security announcements, posters, newsletters, short videos, and feedback about previous phishing attacks (Educause, 2020). These kinds of brief materials can be used to cover a large number of employees and can be easily embedded in individuals' day-to-day work. However, an information technology (IT) department reportedly requires a large amount of resources to analyze, design, develop, and administer such phishing awareness programs (Osterman Research, 2019). Thus, to facilitate antiphishing behavior and performance in organizations, researchers and practitioners should be able to systematically analyze the effectiveness of these educational programs under various conditions.

Several studies have examined how antiphishing training helps individuals differentiate between legitimate and phishing emails (Jampen et al., 2020; Jensen et al., 2017; Kumaraguru et al., 2007a; Schuetz et al., 2020; Sheng et al., 2007; Silic & Lowry, 2020; Wen et al., 2019). In particular, many of these studies often use feedback to improve the learning outcomes of antiphishing training. For example, Kumaraguru et al. (2007a) found that trainees learn more effectively when they get immediate feedback provided through embedded training after the trainee fall for the phishing attack than when they receive the feedback later via email. The game-based training provides both evaluative (e.g., correct or incorrect answers) and informative (e.g., why it is phishing) feedback. Also, Jensen et al. (2017) studied the relative effectiveness of training techniques and showed that mindfulness feedback is superior to rule-based feedback in helping people to thwart phishing attacks. Wen et al. (2019) examined a role-play antiphishing game and stated providing feedback as one of their primary design principles. Silic and

Lowry (2020) showed that a gamified security training program which continues to provide feedback is more effective than a simple email-communication security training program. Jampen et al. (2020) reviewed the effectiveness of different antiphishing training programs and confirmed the long-term effect of providing feedback. Feedback not only reports actual performance but also informs recipients how antiphishing techniques could be used in similar situations (Carver & Scheier, 1982; Lam et al., 2011; Earley et al., 1990; Rakoczy et al., 2008). Thus, proper feedback in antiphishing training can sharpen individuals' awareness of the pervasiveness of phishing and how to detect it.

Although prior research revealed important aspects of antiphishing training, our knowledge in this research topic is lacking in several critical areas. First, phishing research has generally focused on detection accuracy as a major behavioral outcome (Jensen et al., 2017; Silic & Lowry, 2020; Wang et al., 2017). However, little is known about a common case in which individuals decide not to respond to an incoming email (Waterloo News, 2019). Decision avoidance, which inevitably entails a loss of potential benefits from interacting with others, is a universal phenomenon under uncertainty (Anderson, 2003). Therefore, to better understand the role of antiphishing training, it is important to examine not only detection accuracy but also decision avoidance. Second, assessing feasibility is an essential step before making decision under uncertainty, and the same is highly likely when individuals are faced with potential phishing emails (Bagozzi et al., 2003b; Dutton & Webster, 1988). However, a search of the literature on phishing shows this behavior has rarely been examined in the context of antiphishing training. To better explain whether individuals correctly identify and respond to potential phishing scams, it requires a careful examination of the role of perceived feasibility. Last, but not least, our understanding is severely limited on how feedback characteristics interact with phishing characteristics in regulating antiphishing behavior and performance. Because each phishing attack is unique, the effectiveness of feedback is unlikely to be identical across phishing scams. A systematic examination into the effect of feedback on antiphishing behavior and performance cannot be complete without an additional analysis of phishing characteristics.

The objective of this study is to develop and test a model on the effect of feedback on decision avoidance and detection accuracy in the context of antiphishing training. In particular, our model provides a theoretical account of how the relationships between phishing characteristics (e.g., phishing cue saliency) and antiphishing outcomes change with feedback characteristics (e.g., feedback quantity). Drawing on theories of feasibility (Bagozzi et al., 2003b; Dutton & Webster, 1988), we proposed that perceived detection feasibility is a key factor that intervenes between antecedents (feedback and phishing characteristics) and outcomes (decision avoidance and detection accuracy). The proposed model further extends the notion of feasibility with the concept of skill acquisition (Anderson, 1982, 1987) to explain how antecedents interact to influence outcomes. To empirically test the proposed model, we performed three experiments with 514 subjects in the United States. Specifically, given the prevalence of phishing scams relying on false links in an email (PhishLabs, 2019; Proofpoint, 2020), this study focused exclusively on those popular phishing techniques associated with false links. Our results indicate that feedback with a focus on technical information is more effective than feedback with abstract information for correctly distinguishing between phishing and legitimate emails. However, feedback with abstract information tends to make individuals more cautious to make decision about an email's legitimacy. In addition, we show that perceived detection feasibility is essential for a better understanding of antiphishing behavior and performance. Most important, this study demonstrates that the impacts of phishing cue saliency on antiphishing outcomes vary significantly with the quantity of feedback.

Our study contributes to information systems (IS) and phishing research in several ways. First, we are among the few researchers who have examined both decision avoidance and detection accuracy within a coherent theoretical framework to reveal the complex nature of antiphishing behavior and performance. Second, we have theoretically proposed and empirically shown that technical feedback outperforms mindful feedback in facilitating antiphishing performance, at least against phishing scams involving fake links. However, mindful feedback can make individuals more conservative to make decision about an email's legitimacy. Thus, it may have the potential to help individuals avoid future phishing messages that use more advanced phishing strategies rather than spoofing links. Third, drawing on theories of feasibility, we highlight the importance of perceived detection feasibility in antiphishing training. Fourth, we have integrated the notion of feasibility with the concept of skill acquisition to offer a theoretical account of how the impact of perceived detection feasibility on detection accuracy changes with types of feedback. Fifth, drawing on skill acquisition theory, we have also shown an interesting interaction effect between feedback quantity and phishing cue saliency in the context of antiphishing training. Overall, this study contributes significantly to IS research by providing a systemic, theory-driven model of how the offensive (phishing characteristics) and defensive (feedback characteristics) sides of phishing interact to regulate antiphishing behavior and performance.

2. Theoretical Background

2.1. Feedback and Phishing Characteristics

2.1.1. Feedback Characteristics

Feedback refers to information about a person's performance of a task. Feedback can take various forms such as progress bars, leaderboards, points, grades, and texts (Kluger et al., 1994; Werbach & Hunter, 2012). Use of feedback as a tool for informing individuals and members of organizations about their performance has been studied extensively (Kluger & DeNisi, 1996; Lam et al., 2011; Tseng et al., 2019). Prior researchers have considered feedback an important factor in increasing individuals' engagement and performance (Bangert-Drowns et al., 1991; Kluger et al., 1994). Also, in several studies, feedback on antiphishing training has been used to improve antiphishing performance (Jensen et al., 2017; Kumaraguru et al., 2007a; Sheng et al., 2007; Shepherd & Archibald, 2017; Silic & Lowry, 2020). Providing feedback is important because it can "serve as a motivational factor" and "correct illusory performance perceptions" (Jung et al., 2010, p. 728). People have an inherent desire to achieve tasks, and feedback can be used as a base for facilitating antiphishing behavior and performance.

However, the effect of feedback on task performance relies largely on its own characteristics (e.g., type of feedback information, feedback quantity and timing) and also the task characteristics (e.g., task complexity) at hand (Hattie & Timperley, 2007; Kluger & DeNisi, 1996). Although much research exists on antiphishing feedback, no one has systematically investigated how antiphishing behavior and performance vary with feedback characteristics. Thus, it is important to examine the effects that feedback characteristics have on decision avoidance and detection accuracy. In particular, this study compares two types of feedback often examined in the antiphishing literature. These two types are feedback using technical concepts (technical feedback) and feedback using mindful concepts (mindful feedback) (Canova et al., 2015b; Jensen et al., 2017; Stockhardt et al., 2016). Whereas technical feedback refers to feedback that instructs trainees about phishing cues and phishing techniques to help them identify a phishing message, mindful feedback refers to feedback that encourages trainees to allocate attention in processing messages (Jensen et al., 2017). Although both types of feedback are known to be effective in antiphishing training, it also is important to examine their relative effectiveness.

We further examined the role of feedback quantity in subsequent antiphishing outcomes. This study defines feedback quantity as the details of information present in feedback as measured by the amount of text and graphics. Prior research has suggested that more feedback leads to more learning and better task performance (Newell, 1976; Salmoni et al., 1984). However, because individuals have limited cognitive resources, feedback quantity may not always lead to better performance (Kluger & DeNisi, 1996; Lam et al., 2011). Thus, it remains unclear how feedback quantity would influence decision avoidance and detection accuracy, especially in the context of antiphishing behavior.

2.1.2. Phishing Characteristics

Prior research has noted that task characteristics play an important role in predicting task performance (Devine & Kozlowski, 1995; Griffin et al., 1981; Mohammed & Harrison, 2013). For example, Oldham et al. (1976) found that job characteristics significantly influence work performance. Also, Devine and Kozlowski (1995) showed that task characteristics moderate the relationship between decision makers' knowledge level and their decision accuracy. In particular, much research emphasizes that task complexity is among the most important attributes of task characteristics in influencing individuals' behavior and performance (Liu & Li, 2012; Xu et al., 2014).

Task complexity is "the result of the attentional, memory, reasoning, and other information demands imposed by the structure of the task" (Robinson, 2001, p. 29). Thus, a higher level of task complexity requires a larger amount of cognitive resources, which tends to adversely affect decision making and task performance (Braarud, 2001). In the context of decision making, for example, Tversky and Kahneman (1981) claimed that the complexity of practical problems of decision tasks (e.g., portfolio selection) would prevent individuals from integrating existing options. Also, Steele-Johnson et al. (2011) found that both objective and subjective task complexity negatively influenced task performance.

In the context of antiphishing training, an antiphishing task involves an activity of differentiating between legitimate and phishing emails. Successful performance of such an antiphishing task may depend on a variety of factors, but as discussed previously, the complexity of the task is arguably among the most important factors affecting performance. Some phishing messages are easy to detect because of simple and obvious cues (e.g., incorrect names, suspicious sender addresses), whereas others are highly deceptive with complex and concealed cues (e.g., a fake link buried in a long message). Thus, as a factor reflecting the complexity of a task, the saliency of these cues is likely to play a critical role in antiphishing behavior and performance (Downs et al., 2006; Sheng et al., 2007). This study defines phishing cue saliency as the degree to which phishing cues are obvious in an email message. Despite the importance of phishing cue saliency in influencing antiphishing behavior and performance, little research has been conducted on this concept in the IS area.

2.2. Theories of Feasibility

Feasibility is an important concept in understanding task performance in achievement-based contexts such as learning and training. Other researchers have concluded that feasibility is an important predictor of task performance and motivation (Bagozzi et al., 2003b; Klein, 1990). For example, Bagozzi et al. (2003b) argued that perceived feasibility influences an individual's self-commitment to accomplish a chosen goal. Perceived feasibility reflects an individual's subjective expectation about how likely he or she can achieve a goal or task (Bagozzi et al., 2003a, 2003b; Dutton & Webster, 1988; Kanfer & Ackerman, 1989). Dutton and Webster (1988) showed that a task perceived as feasible to resolve attracts more interest from decision makers. Also, Fitzsimmons and Douglas (2011) found that perceived feasibility positively influences entrepreneurial intentions. In short, theories of feasibility suggest that perceived feasibility determines individuals' decision making and performance.

In achievement-based contexts, perceived feasibility is known to be critical in regulating task performance at the individual level (Bagozzi et al., 2003b; Klein, 1990). Likewise, in antiphishing training, perceived feasibility is likely to be a significant determinant of decision making and performance. In this study, perceived detection feasibility refers to as an individual's subjective probability of detecting a phishing message. Despite the importance of perceived detection feasibility in antiphishing training, our understanding of this feasibility concept is quite limited. It remains to be seen whether this factor would be significant in determining antiphishing behavior and performance even after controlling for already well-established factors.

2.3. Skill Acquisition Theory

Anderson (1982) proposed a framework for skill acquisition that includes three knowledge phases in the development of a cognitive skills: (1) declarative knowledge, (2) knowledge compilation, and (3) procedural knowledge (Anderson, 1982, 2010). Declarative knowledge refers to knowledge about facts (Anderson, 2010). This phase of skill acquisition contains all the necessary memory and reasoning processes that allow an individual to obtain an understanding of task requirements (Anderson, 2010; Kanfer & Ackerman, 1989). The second phase, knowledge compilation, is "the process by which the skill transits from the declarative stage to procedural stage" (Anderson, 1982, p. 369). During this phase, the best course of action sequences is chosen out of numerous alternative sequences that can be used to accomplish the same task requirements. As a result, knowledge composition integrates the sequences of procedures required to perform a task into a single procedure. Finally, procedural knowledge is defined as "knowledge about how to perform the task" (Anderson, 2010, p. 205). This final phase of skill acquisition is reached when an individual fundamentally automizes the skill and the task can be accurately performed with little attention (Kanfer & Ackerman, 1989). Complete procedural knowledge is generally achieved after a considerable amount of consistent practice.

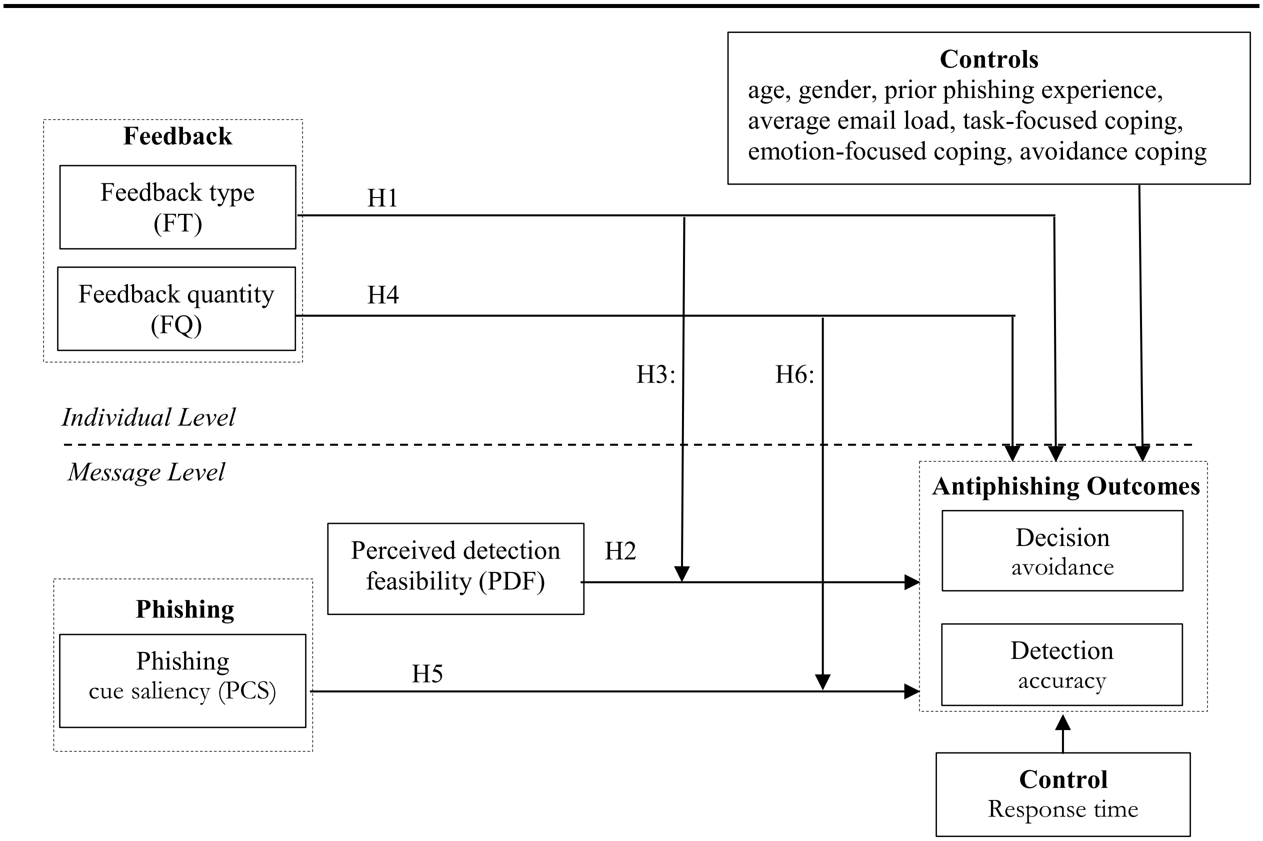
Prior research on skill acquisition noted that declarative knowledge and procedural knowledge are on a continuum; that is, declarative knowledge can be developed into procedural knowledge via knowledge compilation (Anderson, 1982, 2010; Bialystok, 1979). Furthermore, obtaining declarative knowledge requires significant cognitive resources in skill acquisition (Phase 1), but after an individual learns skills through knowledge compilation (Phase 2) and knowledge proceduralization (Phase 3), the demands on cognitive resources are significantly decreased. Accordingly, after knowledge proceduralization is complete, the task can be performed with few attentional resources (Kanfer & Ackerman, 1989; Taatgen et al., 2007).

When organizational interventions (e.g., feedback) and multiple tasks (e.g., phishing quizzes) are provided in training, organizations should consider how trainees can effectively acquire cognitive skills. This is because trainees have a limited cognitive capacity for acquiring new information (i.e., declarative knowledge) and integrating the information and procedures (i.e., knowledge compilation), and applying the combination to different tasks (i.e., procedural knowledge) (Kanfer & Ackerman, 1989; Lam et al., 2011; Tseng et al., 2019). As a result, skill acquisition theory is expected to help better understand antiphishing behavior and performance by taking into account how individuals can effectively acquire and use cognitive skills in performing antiphishing tasks.

3. Research Model and Hypotheses

Figure 8 presents a conceptual model and the research hypotheses proposed in this study to explain decision avoidance and detection accuracy. Hypothesis development is provided below.

Figure 8. Research Model


Controls
age, gender, prior phishing experience, average email load, task-focused coping, emotion-focused coping, avoidance coping
Feedback
Feedback type (FT)
Feedback quantity (FQ)
H1
H4
H3:
H6:
Individual Level
Message Level
Antiphishing Outcomes
Decision avoidance
Detection accuracy
Perceived detection feasibility (PDF)
H2
Phishing
Phishing cue saliency (PCS)
H5
Control
Response time

3.1. Feedback Type and Perceived Detection Feasibility

From the skill acquisition perspective, feedback helps individuals obtain knowledge and correct existing knowledge (declarative knowledge), integrate various procedural knowledge into one simplified action sequence (knowledge compilation), and apply the obtained knowledge to task performance (knowledge proceduralization). Especially, when feedback is provided with direct cues, such feedback is known to further reduce cognitive effort, improve cognitive attention, and eventually lead to higher performance (Kanfer & Ackerman, 1989). This happens because feedback with direct cues helps people spend their cognitive resources only on the essential action steps as specified in such feedback (Jung et al., 2010;

Roch et al., 2000). Accordingly, technical feedback, which focuses on technical attributes manipulated by the attackers, is likely to facilitate antiphishing behavior and performance. In particular, technical feedback—which contains technical terms such as “HTTPs,” “website address,” and “email domain” as well as specific guidelines such as “hover your mouse on the link”—makes it easy for people to apply specific antiphishing tips to the task of detecting phishing messages.

Unlike technical feedback, mindful feedback encourages individuals to dynamically assign cognitive attention and improve awareness of context. In doing so, mindful feedback typically includes speculative suggestions such as “stop” and “don't mindlessly act on an email” (Jensen et al., 2017). Mindful feedback can be helpful in making individuals conscious of checking incoming emails, however, mindful feedback is unlikely to be the most effective form for directing an individual’s attention to specific technical cues present in phishing emails. Mindful feedback focuses on what to do rather than how to do; thus, individuals provided with mindful feedback may not have a clear idea of how mindful principles can be applied to the specific task of antiphishing. In summary, individuals who are given technical feedback can acquire procedural knowledge more easily than those instructed in mindful feedback. Furthermore, when trainees successfully compile and proceduralize declarative knowledge, they can effectively perform a task with confidence and little cognitive burden (Kanfer & Ackerman, 1989). Accordingly, compared with mindful feedback, technical feedback is more effective in helping individuals correctly recognize phishing messages. Mindful feedback, on the other hand, will make individuals more cautious to make decision about whether an email is phishing or legitimate. Thus, we hypothesize that

H1a: Technical feedback leads to less decision avoidance than mindful feedback.

H1b: Technical feedback leads to higher detection accuracy than mindful feedback.

In the context of antiphishing training, detection feasibility is an individual’s subjective evaluation of the likelihood of detecting a phishing message. Other researchers (e.g., Bagozzi et al., 2003a; Fitzsimmons & Douglas, 2011) have noted that feasibility is aligned with expectancy, a belief that an individual’s effort will result in accomplishment of a desired goal (Vroom, 1964). When their expectancy is high, individuals are motivated to spend their cognitive resources, leading to better task performance (Bagozzi et al., 2003a; Vroom, 1964; Kanfer & Ackerman, 1989). In the context of antiphishing training, when individuals believe that they can identify a phishing email, they are likely to take actions to complete a phishing task. Moreover, people high on perceived detection feasibility will expend more cognitive effort to meet their expectation of identifying a phishing email. All things being equal, such an increased effort toward an antiphishing task will increase detection accuracy. From the above reasoning, we hypothesize that

H2a: Perceived detection feasibility negatively influences decision avoidance.

H2b: Perceived detection feasibility positively influences detection accuracy.

Feasibility reflects “the amount of time and effort one has to invest” and “how aspects of an action” to achieve a certain task (Liberman & Trope, 1998, p. 7). Thus, the effect of perceived feasibility on detection accuracy will be reinforced when trainees have sufficient cognitive resources and knowledge on how to perform phishing tasks. As previously noted, technical feedback provides actual antiphishing tips and techniques, and thus individuals can easily understand and apply such feedback to the phishing task at hand. In skill acquisition, individuals given technical feedback can easily compile and proceduralize their phishing knowledge and can allocate their cognitive resources efficiently. Accordingly, technical feedback, which facilitates procedural knowledge, is expected to enhance the successful realization of mere perceived feasibility. Thus, under the condition of technical feedback, the effect of perceived detection feasibility on detection accuracy will be strengthened.

In contrast, individuals receiving mindful feedback tend to face more difficulty in proceduralizing antiphishing knowledge. Because they do not receive detailed instruction about how to apply specific antiphishing tips to detect phishing messages. Mindful feedback has less power to reinforce the successful realization of one’s feasible ideas. Thus, under mindful feedback, the effect of perceived feasibility on detection accuracy will be weaker. Overall, we hypothesize that the effect of perceived detection feasibility on detection accuracy will be stronger under the technical feedback condition than under the mindful feedback condition.[7]

> *H3: There will be an interaction effect between feedback type and perceived detection feasibility on detection accuracy in such a way that the positive effect of perceived detection feasibility will be stronger under the technical feedback condition than under the mindful feedback condition.*

3.2. Feedback Quantity and Phishing Cue Saliency

We hypothesized earlier about how technical feedback and mindful feedback will have different effects on individuals' decision avoidance and detection accuracy. We further proposed that decision avoidance and detection accuracy would vary with the amount of information provided in technical feedback. The quantity of technical feedback is defined as the details of information about phishing cues and antiphishing tips present in an educational message. Prior IS research has suggested a sufficient amount of information is an important attribute of information quality that leads to IS success (Kim et al., 2004; Palmer, 2002). Likewise, feedback research has shown that more feedback enables individuals to make better use of feedback information to learn important task strategies and increase task performance (Bilodeau, 1966; Cook, 1968; Komaki et al., 1980). As discussed previously, individuals who are provided with educational information in a form conducive to procedural knowledge (e.g., technical feedback) are expected to perform conventional antiphishing tasks more effectively than those provided with abstract, high-level information (e.g., mindful feedback) (Anderson, 1982, 2010). Furthermore, individuals given a high quantity of technical feedback will be able to acquire more information about phishing emails and how to detect them. Subsequently, those provided with ample accessible procedural knowledge that can be readily applied to phishing tasks are likelier to be capable of differentiating between legitimate and phishing emails. Accordingly, we hypothesize that the quantity of technical feedback leads to reduced decision avoidance and increased detection accuracy.

---

[7] Unlike detection accuracy, which has a relatively solid objective criterion, however, decision avoidance is an outcome largely controlled by a person's own willingness and subjective evaluations of the self and environment; thus, the impact of perceived detection feasibility, which also is a subjective factor, on decision avoidance seems unvarying across different situations. Thus, we do not hypothesize an interaction effect between feedback type and perceived detection feasibility on decision avoidance.

H4a: Quantity of technical feedback negatively influences decision avoidance.

H4b: Quantity of technical feedback positively influences detection accuracy.

Complex tasks require more cognitive resources because they impose increased stress and cognitive loads on people (Lam et al., 2011; Norman & Bobrow, 1975; Robinson, 2001; Wickens, 1984). For example, in the initial stage of skill acquisition, Kanfer and Ackerman (1989) noted that complex tasks divert individuals' cognitive resources toward self-regulatory processes and away from task-focused learning, leading to inefficiency in task performance. In the context of antiphishing, a phishing email with high cue saliency, by definition, involves numerous phishing cues. Accordingly, such a phishing scam can be detected with minimal cognitive demand because its numerous cues make it readily identifiable. In contrast, a phishing email with low cue saliency has few phishing cues and consequently is harder to recognize; identifying such a phishing attempt requires more mental exertion, which inevitably leads to inefficiency in antiphishing performance (Kanfer & Ackerman, 1989). In general, the discussion mentioned previously leads us to expect that phishing cue saliency will be positively associated with detection accuracy.[8]

H5: Phishing cue saliency positively influences detection accuracy.

We earlier predicted that phishing cue saliency would have a significant impact on detection accuracy. We further proposed that feedback quantity would moderate the effect of phishing cue saliency on antiphishing performance. If individuals are given a small amount of technical feedback (low quantity feedback), the level of cue saliency in an email is not a major issue for antiphishing performance. This is because when individuals have limited knowledge on how to detect phishing cues, they are likely to focus only on the key signal they are trained to focus on to the exclusion of other potentially important cues. Put differently, for those with a minimal set of antiphishing tips, it is cognitively demanding to identify peripheral phishing cues that currently lie outside the realm of procedural knowledge in their mental representation (Kanfer & Ackerman, 1989; Lam et al., 2011).

---

[8] We do not hypothesize the effect of phishing cue saliency on decision avoidance. This is because decision avoidance is mainly a function of an individual's subjective evaluation of the task at hand (i.e., perceived detection feasibility) but not of objective task complexity (i.e., phishing cue saliency).

In contrast, when a large amount of technical feedback is given (high quantity feedback), this type of information is expected to facilitate the accumulation of antiphishing tips in the form of procedural knowledge. Such abundant tips are likelier to correspond to some of the phishing cues present in a phishing email. Accordingly, when individuals have plenty of knowledge on how to detect phishing cues, the level of cue saliency in an email will exert a larger impact on detection accuracy. In this condition of high quantity feedback, even peripheral cues can be identified more easily with little cognitive overload (Lam et al., 2010; Tseng et al., 2019). As a result, the effects of phishing cue saliency on detection accuracy will be stronger with high quantity information than with low quantity information.[9] Thus, we hypothesize that

> *H6: There will be an interaction effect between quantity of technical feedback and phishing cue saliency on detection accuracy in such a way that the positive effect of phishing cue saliency will be stronger under the high feedback quantity condition compared with what occurs under the low feedback quantity condition.*

4. Methods

To test our research hypotheses, we conducted three experiments as summarized in Table 9. Experiment 1 examined the effect of feedback type (technical vs. mindful) on decision avoidance and detection accuracy (H1). In Experiment 2, we tested the effect of perceived detection feasibility (H2) and its interaction with feedback type (H3). Experiment 3 further investigated the effects of feedback quantity (H4) and phishing cue saliency (H5) as well as their interaction effect (H6).

We conducted Experiment 1 and 3 using Amazon Mechanical Turk workers, while we conducted Experiment 2 on a US-based online panel. We created web-based surveys to conduct our experiments. In general, for each experiment, we first introduced phishing emails as malicious emails trying to steal personal information and invited the participants to take part in interesting phishing quizzes. Afterward, our three experiments followed similar procedures as shown in Figure 9. All participants needed to complete a preliminary phishing test, including email quizzes. The participants were required to read two emails and indicate their judgments about each email (phishing; legitimate; skip the question). Next, we showed different feedback materials to different experimental groups and then conducted a manipulation check. After reading the feedback, the participants were asked to complete a main phishing test including a series of new email quizzes. Finally, the participants filled out a survey questionnaire for us to measure several control variables. The details of the methods and results will be described in each experiment. Appendix C presents all the measurement items, and Appendix D shows the examples of feedback, phishing emails, and legitimate emails in the experiments.

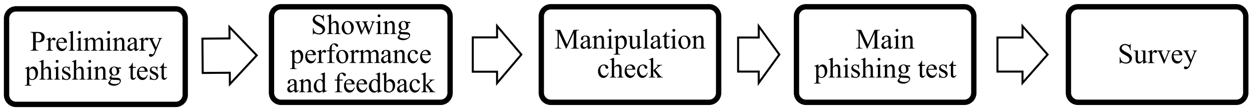
---

[9] As mentioned earlier, decision avoidance is a function of perceived detection feasibility. Thus, we do not expect an interaction effect between feedback quantity and phishing cue saliency on decision avoidance.

Table 9. Summary of Experiments

		Experiment 1	Experiment 2	Experiment 3
Hypothesis		H1	H2, H3	H4, H5, H6
Design		Single-factor between-subjects design	Single-factor design	2 × 2 mixed deign
Subject		MTurk workers (n = 130)	Online panel (n = 110)	MTurk workers (n = 274)
Dependent Variable		• Decision avoidance (0: decision making, 1: decision avoidance) • Detection accuracy (-1: incorrect, 0: skipping, 1: correct)		
Individual Level IV	Manipulated	• Feedback type (technical vs. mindful)	• Feedback type (technical vs. mindful)	• Feedback quantity (low vs. high)
	Measured	Age, gender, prior phishing experience, average email load, preliminary training detection accuracy, task-focused coping, emotion-focused coping, avoidance coping		
Message Level IV	Manipulated	NA	NA	• Phishing cue saliency (low vs. high)
	Repeatedly measured	NA	• Perceived detection feasibility • Response time	• Perceived detection feasibility • Response time
Analysis Methods		• ANCOVA	• Generalized estimating equation	• Generalized estimating equation

Figure 9. Experimental Procedure



4.1. Experiment 1

Experiment 1 tested the relative effectiveness between technical feedback and mindful feedback (H1). This experiment employed a single factor (feedback type: technical vs. mindful) between-subjects design.

4.1.1. Treatment

Two survey websites were customized for our manipulation of feedback type (technical vs. mindful), and the participants were randomly assigned to one of the two survey websites.

Specifically, as mentioned before, participants were asked to complete a preliminary phishing test, including one legitimate email and one phishing email. The participants were required to indicate their judgments about each email (phishing; legitimate; skip the question). After the preliminary test, participants in each group were shown different feedback. For technical feedback, we presented the prior phishing email with the fake URL address next to its fake link and explicitly taught the participants how to detect the fake link within this email. For mindful feedback, we showed participants feedback adapted from the mindfulness training material (Jensen et al., 2017). It reminded participants to be cautious with email links and to think more about the email's requirements before acting on an email. Appendix D shows the feedback materials in the test.

4.1.2. Sample

We collected data on Amazon Mechanical Turk where we could reach out to a large number of potential participants with diverse backgrounds. The workers could participate in the study if they were at least 18 years old and were located in the United States. In order to ensure the data quality, we followed a rigorous data cleaning procedure. We removed responses with the same IP address, location or the same birthdate with similar start times to make sure that the participants took the survey only once and independently. We also deleted the participants who didn't complete the entire experiment and those who didn't provide valid answers to the survey questions (e.g., invalid birthdate). We also checked the locations based on IP address and Location Longitude/Latitude. The participants who were not located in the United States were also excluded. The final sample consisted of 130 participants. The average age of the participants was 35.63, and 47.69% were female[10].

4.1.3. Experimental Procedures

At the beginning of the experiment, we first introduced phishing emails as malicious emails trying to steal personal information and invited the participants to take part in interesting phishing quizzes. After the introduction, participants were required to complete a preliminary phishing test, including one legitimate email and one phishing email in random order. As mentioned before, the participants were required to indicate their judgments about each email (phishing; legitimate; skip the question). Afterward, we presented the participants with one of the two types of feedback (i.e., technical or mindful).

After reading the feedback, all participants completed the main phishing test, including six new email quizzes. The participants were still asked to report their decisions about each email's legitimacy. To keep participants' attention, we also informed them that they would gain $1 as a base rate, earn extra $0.05 for each correct answer, lose extra $0.05 for each wrong answer, and get no extra incentive for skipping the answer in the main test. Specifically, all participants were exposed to three legitimate emails and three phishing emails in random order. Finally, to measure our control variables, we asked participants about their coping responses (task-focused, emotion-focused, and avoidance) as well as demographic information.

4.1.4. Measures

The dependent variables, decision avoidance and detection accuracy, were measured based on participants' objective performance in the phishing test. To measure decision avoidance, we coded a participant's answer to each email quiz as 1 if the person chose "skip the question" and 0 if the person made decision. In addition, we coded a wrong answer to each quiz as -1, a correct answer as 1 and a skipping answer as 0 to measure detection accuracy.

---

[10] We used three categories ("Female", "Male" or "Other") to record gender. However, there were only two participants who chose the "Other" category and we thus didn't include those participants in the final dataset. The other two experiments also followed this procedure.

We included participants' coping responses as control variables because Wang et al. (2017) showed that these responses have direct impacts on antiphishing performance in a similar phishing test. Based on Wang et al. (2017), we measured three types of coping responses—task-focused, emotion-focused, and avoidance. Finally, as controls, we also measured demographic information such as participants' prior phishing experience, average email load, age, and gender.

4.1.5. Results and Discussion

We analyzed the impact of feedback type on decision avoidance and detection accuracy using an analysis of covariance (ANCOVA). Table 10 shows the results of the ANCOVA. For ANCOVA, we aggregated the scores of the six email quizzes in the main test for decision avoidance and detection accuracy, respectively. Thus, decision avoidance ranged from 0 to 6 and detection accuracy ranged from – 6 to + 6.

Table 10. Experiment 1: ANCOVA Results for Decision Avoidance and Detection Accuracy

	Decision Avoidance				Detection Accuracy			
	Mean Square	F	*p*-value	Effect Size	Mean Square	F	*p*-value	Effect Size
Manipulation								
FT (H1)	2.99	4.28	0.041	0.034	244.41	36.34	< .001	0.232
Control Variables								
Age	0.05	0.07	0.799	0.001	1.77	0.26	0.609	0.002
GEN	0.12	0.18	0.674	0.001	0.14	0.02	0.885	0.000
PPE	0.26	0.38	0.541	0.003	0.20	0.03	0.862	0.000
AEL	0.29	0.42	0.518	0.003	34.91	5.19	0.024	0.041
PTDA	0.74	1.06	0.306	0.009	14.19	2.11	0.149	0.017
TC	1.14	1.63	0.204	0.013	21.59	3.21	0.076	0.026
EC	0.36	0.52	0.472	0.004	4.80	0.71	0.400	0.006
AC	1.06	1.52	0.220	0.013	3.24	0.48	0.489	0.004
Model Fit								
R^2				0.066				0.306
Adjusted R^2				-0.004				0.254

Notes:

- FT = feedback type (mindful = 0, technical = 1); GEN = gender (male = 0, female = 1); PPE = prior phishing experience; AEL = average email load; PTDA = preliminary training detection accuracy; TC = task-focused coping; EC = emotion-focused coping; AC = avoidance coping.

After we controlled for other variables, the results showed a significant difference in the mean decision avoidance between the technical feedback group and the mindful feedback group (F (1, 120) =

4.28, $p < 0.05$). By comparing the estimated marginal means, we found that participants in the technical feedback group showed less decision avoidance ($M = 0.15$) than those in the mindful feedback group ($M = 0.46$), suggesting support for H1a. In addition, a significant difference occurred in the mean detection accuracy between the technical feedback group and the mindful feedback group (F (1, 120) = 36.34, $p < 0.001$) after controlling for control variables. By comparing the estimated marginal means, we found that the technical feedback group had higher detection accuracy ($M = 4.10$) than the mindful feedback group ($M = 1.28$), suggesting support for H1b.

4.2. Experiment 2

Experiment 2 aimed at testing the impact of perceived detection feasibility on decision avoidance and detection accuracy (H2a and H2b), as well as the interaction effect between feedback type (mindful vs. technical) and perceived detection feasibility on detection accuracy (H3).

4.2.1. Treatment

The feedback materials in Experiment 2 were consistent with those in Experiment 1. We still created two survey websites for the manipulation of feedback type (technical vs. mindful), and randomly assigned the participants to one of the two survey websites.

All participants were exposed to two phishing emails in random order in a preliminary test, and the participants were required to indicate their judgements about each email. After each quiz, we provided the answer (i.e., if their judgments about the email were correct) to the participants, followed by the type of feedback they were assigned to. For the mindful feedback, we used the same feedback material as in Experiment 1. For the technical feedback, we showed participants the image of each phishing email and showed them how to detect the fake link within each phishing email. The feedback design was consistent with the one we used in Experiment 1, except that we used different phishing emails.

4.2.2. Sample

We conducted Experiment 2 by inviting potential participants via email from a US-based online panel. Similar to Experiment 1, we still excluded responses with the same IP address, location or the same birthdate with similar start times. We also removed the participants who didn't complete the experiment and those who provided invalid answers to the survey questions. After the data cleaning procedure, the final sample consisted of 110 participants. The average age was 46.44 and 47.27% were female.

4.2.3. Experimental Procedures

We still invited the participants to take part in interesting phishing quizzes. The participants then needed to take a preliminary test including two phishing emails and indicate their judgments about each email. Afterward, they received both the answer to each quiz and the feedback material they were assigned to.

After they read the feedback, all participants were asked to complete the main test, including four new email quizzes. In order to focus their attention, we also informed them that they would get a base rate ($1.3), in addition, they would earn extra $0.05 for each correct answer, lose extra $0.05 for each wrong answer, and get no extra incentive for skipping the answer. The quizzes were chosen from the quizzes we used in Experiment 1. Specifically, all participants were exposed to three legitimate emails and one phishing email in random order. Except from answering each quiz, they were also asked to report their perceived detection feasibility toward each email quiz. At last, we asked participants about their coping responses and demographic information.

4.2.4. Measures

We measured decision avoidance, detection accuracy, coping responses, and other demographic variables by using the same method as in Experiment 1. Additionally, we measured perceived detection feasibility for each email quiz in the main phishing test by using four items such as "It is possible to determine whether the email is phishing" and "It is feasible to determine whether the email is phishing" (Bagozzi et al., 2003b; Dutton & Webster, 1988). We also recorded the time each participant spent on each quiz in the main test by using the survey website as a message-level control variable.

4.2.5. Construct Validation

We performed a confirmatory factor analysis to examine the quality of our measurement items. Appendix E presents the descriptive statistics for all research variables and the results of construct validation. We examined composite reliability, and the scores for all measures exceeded the cutoff value of 0.70, indicating satisfactory reliability (Bagozzi & Yi, 1988). We tested convergent validity through item loadings and average variance extracted (AVE). All measurement items loaded significantly on the assigned construct ($p < 0.001$), and the scores of AVEs were all above the cutoff value of 0.50 (Fornell & Larcker, 1981). All the measures showed satisfactory convergent validity. We also used chi-square tests to examine discriminant validity by comparing a series of model pairs (Segars & Grover, 1998). The chi-square difference tests were all significant, suggesting the satisfactory discriminant validity of our measures. We also compared the AVE for each construct to its correlations with other constructs (Gefen & Straub, 2000). For each construct, the AVE was greater than its correlations with other constructs, indicating discriminant validity.

4.2.6. Generalized Estimating Equations

Our data contained repeated measures of dichotomous outcomes (decision avoidance) and ordinal outcomes (detection accuracy) nested within individuals, along with repeated measures of perceived detection feasibility, response time, and individual-level control variables. Generalized estimating equations (GEE) provided the means to analyze the repeated ordinal/dichotomous outcomes nested within individuals.

We analyzed the two dependent variables separately. Table 11 presents the results of GEE for decision avoidance and detection accuracy, respectively. For each dependent variable, we built five alternative models. Model 1 included only the individual-level treatment variable (i.e., feedback type). Model 2 included the treatment variable and the message-level research variable (i.e., perceived detection feasibility). Model 3 included coping responses as control variables. Model 4 included all other control variables. Model 5 included the interaction term between feedback type and perceived detection feasibility.

For decision avoidance, we calculated model fit indices such as Quasi Likelihood under Independence Model Criterion (QIC) and Corrected Quasi Likelihood under Independence Model Criterion (QICC) to compare the model fit. As shown in Model 1 to Model 3, the model fitted better as

Table 11. Experiment 2: GEE Results for Decision Avoidance and Detection Accuracy

	Decision Avoidance					Detection Accuracy				
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 1	Model 2	Model 3	Model 4	Model 5
Intercept	-4.33^{***}	-4.55^{***}	-6.86^{***}	-8.90^{***}	-9.27^{***}					
Threshold										
-1.00|0.00						-0.89^{***}	-0.91^{***}	-0.92^{***}	-0.94^{***}	-0.94^{***}
0.00|1.00						-0.80^{***}	-0.81^{***}	-0.82^{***}	-0.84^{***}	-0.84^{***}
Level 2										
Manipulation										
FT	$-2.11^{\dagger}$	-1.74	$-2.03^{\dagger}$	-2.57^{*}	-3.58^{*}	-0.71^{**}	-0.69^{**}	-0.65^{**}	-0.65^{**}	-0.71^{**}
Individual Factors										
Age				0.76	0.76				-0.04	-0.05
GEN				3.26	3.18				0.24	0.22
PPE				1.01	0.94				0.16	0.14
AEL				-0.37	-0.35				0.01	0.01
PTDA				-0.57	-0.57				0.02	0.01
TC			0.97	1.19	1.19			0.12	0.10	0.10
EC			0.56	0.45	0.44			0.00	0.01	-0.02
AC			1.58^{***}	2.11^{***}	2.09^{***}			$0.19^{\dagger}$	0.14	0.13
Level 1										
PDF (H2)		-0.95^{*}	-1.94^{***}	-2.67^{***}	-2.87^{***}		-0.23^{*}	-0.27^{*}	-0.28^{*}	-0.31^{*}
RT				0.14	0.10				0.03	0.03
Interaction										
FT * PDF (H3)					-0.56					$-0.41^{\dagger}$
Model Fit										
QIC	85.50	83.52	72.40	153.70	155.10					
QICC	85.60	83.61	72.70	154.60	156.10					
Wald Statistics						-	3.92^{*}	3.91	2.55	$3.01^{\dagger}$
DF						-	1	3	6	1

Notes:

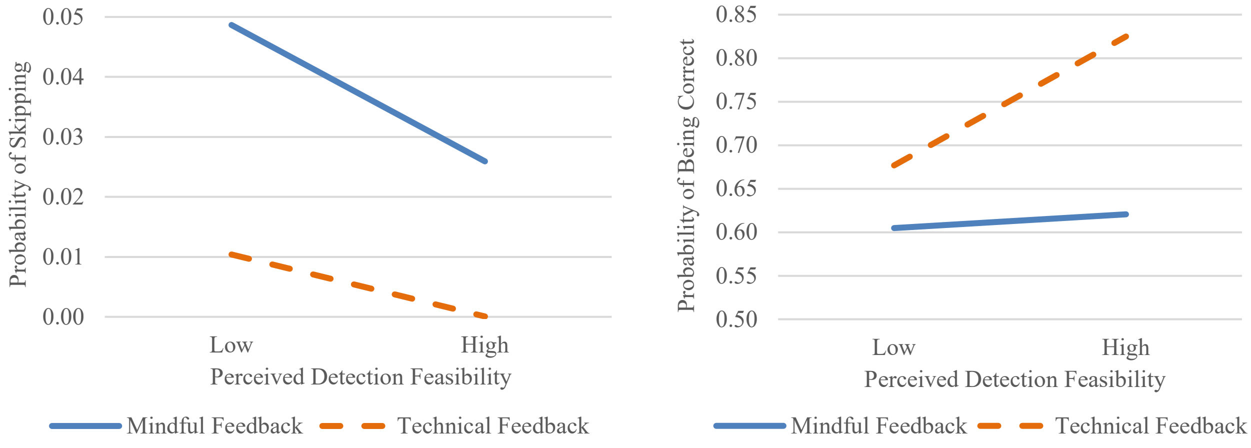
- Individual level: N = 110; Message level: N = 440
- $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\dagger}p < 0.10$ (two-tailed).
- For model fit, QIC and QICC were used for decision avoidance and Wald statistics were used for detection accuracy.
- For GEE with ordinal outcomes, a negative (positive) sign indicates a positive (negative) effect on the outcome variable.
- FT = feedback type (mindful = 0, technical = 1); GEN = gender (male = 0, female =1); PPE = prior phishing experience; AEL = average email load; PTDA = preliminary training detection accuracy; TC = task-focused coping; EC = emotion-focused coping; AC = avoidance coping; PDF = perceived detection feasibility; RT = response time to each quiz.

we added perceived detection feasibility and coping responses. However, except from coping responses, other control variables didn't significantly affect decision avoidance and they didn't increase the model fit. However, we still kept all the control variables to explicitly partial out their effects. The results in Model 4 showed that perceived detection feasibility was significantly and negatively associated with decision avoidance (-2.67, $p < 0.001$), supporting H2a. We didn't hypothesize that there was an interaction effect between feedback type and perceived detection feasibility, and Model 5 also showed that the interaction effect between feedback type and perceived detection feasibility was insignificant.

For detection accuracy, Wald tests were performed to compare the nested models consecutively.[11] As presented in Table 11, adding the control variables didn't increase the model fit but we kept the control variables due to the same reason as mentioned before. We chose Model 5 as the final model for hypotheses testing. As shown in Model 5, perceived detection feasibility was significantly and positively associated with detection accuracy (-0.31, $p < 0.05$), supporting H2b. The interaction effect between feedback type and perceived detection feasibility was marginally significant (-0.41, $p = 0.08$), and thus H3 was only marginally supported.

We also depicted in Figure 10 the interaction plots for decision avoidance and detection accuracy. For decision avoidance, the change in the probability of skipping a quiz from low to high perceived detection feasibility doesn't differ between the technical and mindful feedback groups. For detection accuracy, as perceived detection feasibility increases, the probability of correctly answering a quiz becomes much higher in the technical feedback group than in the mindful feedback group. In summary, we can conclude that perceived detection feasibility is negatively associated with decision avoidance and

Figure 10. Experiment 2: Interaction between Feedback Type and Perceived Detection Feasibility



A: Decision Avoidance B: Detection Accuracy

Notes:
- Perceived detection feasibility is categorized as high (low) when the value is above (below) the mean.
- The probability of skipping and the probability of being correct were estimated based on the results of GEE.

---

[11] QIC and QICC model fit indices cannot be calculated for GEE with ordinal outcomes.

positively associated with detection accuracy. Moreover, we found a marginally significant interaction effect between feedback type and perceived detection feasibility.

4.3. Experiment 3

We conducted Experiment 3 to test H4, H5, and H6. We used a 2 by 2 mixed design with feedback quantity (low vs. high) as the between-subjects factor and phishing cue saliency (low vs. high) as the within-subjects factor to study how people's decision avoidance and detection accuracy for various emails would differ when receiving different levels of feedback quantity.

4.3.1. Treatment

For feedback quantity, as in the prior experiments, participants had to complete a preliminary phishing test that included two email quizzes in random order. After the preliminary test, participants were randomly assigned to the two levels of feedback quantity. In both the low and high feedback quantity group, we used a real phishing email different from any used in the previous test to inform participants of some phishing cues and antiphishing techniques. Specifically, for low feedback quantity, we presented the phishing email and showed only how to detect the fake link within this email. We used the same phishing email for high feedback quantity, but this time we informed the participants of several phishing cues such as the fake link, the fake email domain, a general greeting and also explained phishing techniques such as using an urgent tone. Appendix D shows the details of the feedback design.

For phishing cue saliency, in the main test, three emails were in the low phishing cue saliency group and the other three were in the high phishing cue saliency group (two phishing emails and one legitimate email in each group). All participants needed to complete the six quizzes. We manipulated phishing cue saliency for phishing emails by varying the number of phishing cues. For a phishing email with low phishing cue saliency, the email was well-crafted from the alleged sender's genuine email. The only phishing cue was the fake link. For an email with high phishing cue saliency, the email contained, in addition to the fake link, a suspicious email domain as well as one of the following cues: an urgent tone, minor grammar errors, or general greetings. The underlying assumption of the latter email was that the more the phishing cues it contained, the likelier people would notice at least some of the cues and recognize the phishing email.

4.3.2. Sample

We conducted Experiment 3 by inviting participants on Amazon Mechanical Turk. The participants needed to be at least 18 years old and located in the United States. We followed the same data cleaning procedure as that in Experiment 1. After the data cleaning procedure, the sample consisted of 295 participants. The average age was 38.17 and 50.51% were female.

4.3.3. Experimental Procedures

As in the prior experiments, we invited participants to take interesting phishing quizzes. Participants needed to complete a preliminary phishing test that included two email quizzes in random order. We also provided them with the answer to each email quiz. Afterward, participants would receive their assigned feedback. We then conducted a manipulation check.

After the manipulation check, all participants were asked to complete the main phishing test of six email quizzes in random order. As mentioned before, three emails were in the low phishing cue saliency group and the others were in the high phishing cue saliency group. Similar to previous experiments, they were told that they would get a base rate ($1.5), and they would gain or lose extra money based on their performance (i.e., earn extra $0.01 for a correct answer, lose $0.01 for a wrong answer and get no extra incentive for skipping the answer). After each quiz, we also asked participants for their perceived detection feasibility. We still recorded the time they spent on each email. At last, we asked participants the same questions to measure the individual-level control variables as in the prior experiments.

4.3.4. Measures

We measured decision avoidance, detection accuracy, perceived detection feasibility, coping responses, response time, and other demographic variables by using the same approach as in Experiment 2.

4.3.5. Manipulation Check

We asked participants about the number of tips (one tip or many tips) in their assigned feedback to check if a difference existed in the perceptions about feedback quantity between the low and high feedback quantity groups. Results of the chi-square test indicated that there was a significant difference between the high quantity group and the low quantity group (χ^2 (1, 295) = 215.03, $p < 0.001$). Participants in the high quantity group was likelier to choose “many tips” than those in the low quantity group. The results suggested our manipulation of feedback quantity was successful. There were 21 participants who failed the manipulation check. We excluded those participants in the following analysis. As a result, the final dataset consisted of 274 participants. The average age was 38.36 and 51.09% were female.

4.3.6. Construct Validation

As we did in Experiment 2, we performed a confirmatory factor analysis, including all the 14 research variables into our model, to examine the quality of our measurement items. Appendix E shows the descriptive statistics for all research variables and the results of construct validation. All the constructs showed satisfactory reliability, convergent validity, and discriminant validity.

4.3.7. Generalized Estimating Equations

We tested H4, H5, and H6 using generalized estimating equations to accommodate the hierarchical nature of our data. We still analyzed the two dependent variables separately. Table 12 shows the results of GEE for decision avoidance and detection accuracy, respectively. We built four alternative models for each dependent variable. Model 1 included the individual- and message-level treatment variables (i.e., feedback quantity and phishing cue saliency). Model 2 included the treatment variables and all control variables. Model 3 included perceived detection feasibility and all the prior variables. Model 4 further included the interaction term between feedback quantity and phishing cue saliency.

We still calculated QIC and QICC for decision avoidance. In Table 12, Model 3 showed the best model fit. The control variables didn’t significantly affect decision avoidance but we still kept them in our model to explicitly partial out their effects. The results of Model 3 showed that feedback quantity did not have a significant main impact on decision avoidance (0.01, p = ns), and thus H4a was not supported. The results of Model 3 also showed that phishing cue saliency didn't affect decision avoidance, meanwhile, perceived detection feasibility was significantly associated with decision avoidance (-1.84, $p < 0.001$). This was consistent with our expectation that decision avoidance was influenced by an individual's subjective evaluation of the task at hand. We didn't hypothesize that there was an interaction effect

Table 12. Experiment 3: GEE Results for Decision Avoidance and Detection Accuracy

	Decision Avoidance				Detection Accuracy			
	Model 1	Model 2	Model 3	Model 4	Model 1	Model 2	Model 3	Model 4
Intercept	-3.89***	-3.94***	-5.61***	-5.66***				
Threshold								
-1.00|0.00					-1.46***	-1.51***	-1.62***	-1.62***
0.00|1.00					-1.34***	-1.38***	-1.48***	-1.49***
Level 2								
Manipulation								
FQ (H4)	0.06	0.09	0.01	0.00	0.36*	0.37*	0.33*	0.23
Individual Factors								
Age		-0.03	-0.08	-0.08		-0.18*	-0.19*	-0.19*
GEN		-0.29	-0.66	-0.69		0.12	0.09	0.08
PPE		0.08	0.09	0.10		0.09	0.08	0.08
AEL		-0.19	-0.27	-0.25		-0.05	-0.04	-0.04
PTDA		-0.18	0.02	0.05		0.01	0.08	0.08
TC		-0.19	0.04	0.05		0.05	0.11	0.11
EC		0.05	-0.41	-0.40		0.33***	0.27**	0.28**
AC		-0.08	0.04	0.04		0.11	0.11	0.11
Level 1								
Manipulation								
PCS (H5)	-0.06	-0.06	0.09	0.10	-0.82***	-0.84***	-0.85***	-0.84***
Message-level Factors								
PDF			-1.84***	-1.85***			-0.42***	-0.43***
RT		0.03	-0.75†	-0.76†		0.06	0.01	0.01
Interaction								
FQ × PCS (H6)				-0.86				-0.59*
Model Fit								
QIC	329.79	343.10	259.30	263.70				
QICC	329.81	343.40	259.60	264.00				
Wald Statistics					-	29.16***	34.74***	5.99*
DF					-	9	1	1

Notes:

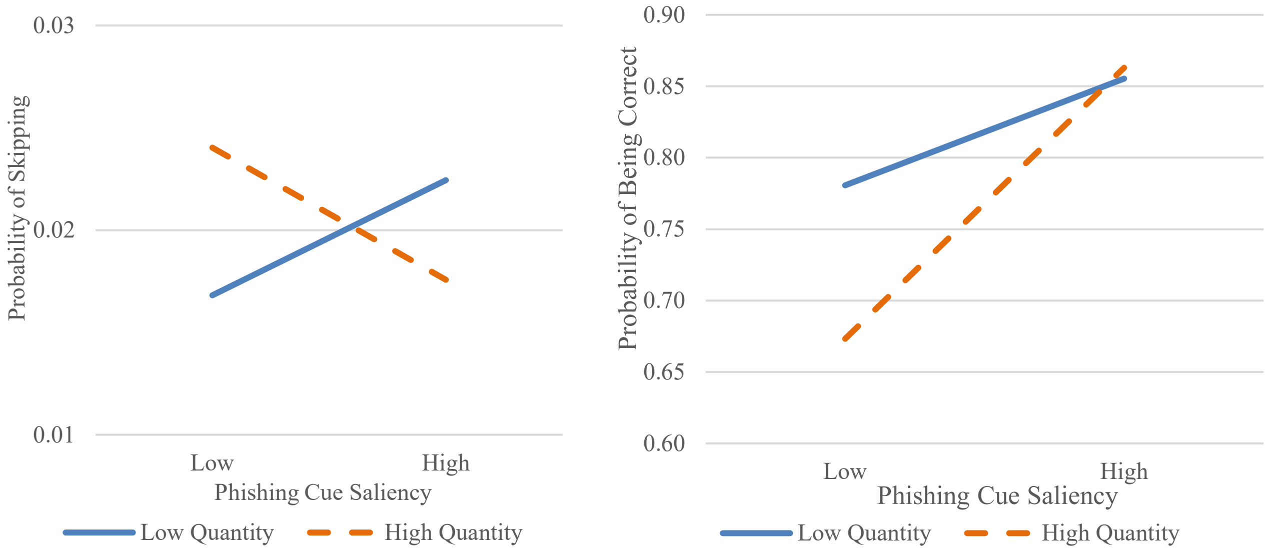
- Individual level: N = 274; Message level: N = 1644
- $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\dagger}p < 0.10$ (two-tailed).
- For model fit, QIC and QICC were used for decision avoidance, and Wald statistics were used for detection accuracy.
- For GEE with ordinal outcomes, a negative (positive) sign indicates a positive (negative) effect on the outcome variable.
- FQ = Feedback quantity (low = 0, high = 1); GEN = gender (male = 0, female = 1); PPE = prior phishing experience; AEL = average email load; PTDA = preliminary training detection accuracy; TC = task-focused coping; EC = emotion-focused coping; AC = avoidance coping; PCS = phishing cue saliency (low = 0, high = 1); PDF = perceived detection feasibility; RT = response time to each quiz.

between feedback quantity and phishing cue saliency on decision avoidance, and Model 4 also showed that the interaction effect was insignificant.

For detection accuracy, as shown in Table 12, we used Wald tests to compare the model fit. The model fit continued to improve as we added more research variables. In Model 3, the effect of feedback quantity just reached a significance level of 0.05 (0.33, $p = 0.048$). However, as shown in Model 4, this effect was replaced by a significant interaction effect between feedback quantity and phishing cue saliency (-0.59, $p = 0.01$). Thus, H4b was not supported. Instead, H6 was supported. The results of Model 4 also indicated a significant positive impact of phishing cue saliency on detection accuracy (-0.84, $p < 0.001$), suggesting support for H5. As mentioned before, feedback quantity did not have a main impact on detection accuracy but moderated the effect of phishing cue saliency. This suggested that the impact of phishing cue saliency was stronger under the high feedback quantity condition than under the low quantity condition.

Figure 11 also presents the interaction plots for decision avoidance and detection accuracy. For decision avoidance, the probability of skipping a quiz is around 2% regardless of the level of feedback quantity or phishing cue saliency. The plot for detection accuracy indicates that as phishing cue saliency changes from low to high, the probability of providing a correct answer increases much more in the high quantity feedback group than in the low quantity feedback group. In addition, for low phishing cue saliency, the probability of giving a correct answer is much higher in the low feedback quantity group; however, for high cue saliency, this probability is a bit higher in the high feedback quantity group. Overall, feedback quantity does not have a main impact on decision avoidance or detection accuracy. Phishing cue saliency has a positive impact on detection accuracy. There is an interaction effect between feedback quantity and phishing cue saliency on detection accuracy.

Figure 11. Experiment 3: Interaction between Feedback Quantity and Phishing Cue Saliency



A: Decision Avoidance B: Detection Accuracy

Notes:
- The probability of skipping and the probability of being correct were estimated based on the results of GEE.

5. Discussion and Conclusion

5.1. Summary of Findings

The objective of this study was to develop and test a model of the role of feedback in antiphishing behavior and performance. Drawing on theories of feasibility, our proposed model highlights perceived detection feasibility as an intervening variable between antecedents (feedback and phishing characteristics) and outcomes (decision avoidance and detection accuracy). Rooted strongly in skill acquisition theory, moreover, it proposes the role of feedback as a moderator between phishing characteristics and antiphishing outcomes. We conducted three experiments to evaluate the proposed model and corresponding research hypotheses. Table 13 shows a summary of research findings from the three experiments.

Table 13. Summary of Research Findings

EXP	Hypothesis	Findings
1	H1	• Technical feedback leads to less decision avoidance and higher detection accuracy than mindful feedback (H1a & H1b supported).
2	H2, H3	• Perceived detection feasibility decreases decision avoidance and increases detection accuracy (H2a & H2b supported). • There is a marginally significant interaction effect between perceived detection feasibility and feedback type on detection accuracy; as expected, however, the interaction is not significant on decision avoidance (H3 marginally supported).
3	H4, H5, H6	• Quantity of technical feedback does not have a significant impact on decision avoidance and detection accuracy (H4a & H4b not supported). • Phishing cue saliency positively influences detection accuracy (H5 supported). • The effect of phishing cue saliency is stronger under the high feedback quantity condition compared with under the low feedback quantity condition (H6 supported).

Using a single-factor design, Experiment 1 investigated the effect of feedback type on decision avoidance and detection accuracy. Our finding based on the data from 130 participants in MTurk provides support for the hypothesis that technical feedback is more helpful than mindful feedback in increasing detection accuracy (H1b). On the other hand, mindful feedback leads to more decision avoidance than technical feedback (H1a). In addition, using a single-factor repeated measures design, Experiment 2 was intended to examine the effects of feedback type and perceived detection feasibility on outcomes. Based on the data collected from 110 members from a nationwide online panel, we found a significant effect of perceived detection feasibility on both decision avoidance and detection accuracy (H2a & H2b). Furthermore, this study suggested that the positive effect of perceived detection feasibility on detection accuracy is stronger for people who receive technical feedback than for people who receive mindful feedback, although this effect was marginally significant (H3). Furthermore, as expected, the type of feedback did not moderate the relationship between perceived detection feasibility and decision avoidance.

Experiment 3 used a 2 × 2 mixed design with 274 participants in MTurk. It examined the effects of feedback quantity and phishing cue saliency on outcomes. Unlike our prediction, the quantity of technical feedback did not have any significant impacts on decision avoidance or detection accuracy (H4a and H4b). A plausible explanation is that the impacts of feedback quantity on outcomes vary with other factors such as types of phishing messages. Meanwhile, we found that phishing cue saliency positively influences detection accuracy (H5). Finally, the effect of phishing cue saliency on detection accuracy was found to be stronger when people receive high feedback quantity than when they receive low feedback quantity (H6).

5.2. Theoretical Contributions

This study makes several theoretical contributes. First, whereas prior research on phishing focused on detection accuracy, decision avoidance has gotten little attention. Our study is among the first to suggest differences in the mechanisms underlying decision avoidance and detection accuracy; in particular, whereas detection accuracy is based on an objective criterion, subjective perceptions largely determine decision avoidance. Furthermore, we found that both have low correlations (-0.23 in Experiment 1, -0.05 in Experiment 2, and -0.10 in Experiment 3). For these reasons, phishing cue saliency influences detection accuracy, but such a relationship could not be found for decision avoidance. Similarly, the impact of perceived feasibility on detection accuracy varies with feedback types, but such an interaction effect was not found for decision avoidance. Thus, it is important for future researchers to examine both decision avoidance and detection accuracy in the contexts of skill acquisition. Overall, this study contributes to IS research by providing an integrated theoretical framework for the different mechanisms regulating detection accuracy and decision avoidance.

Second, drawing on skill acquisition theory (Anderson, 1982), we argue that technical feedback is more effective than mindful feedback in transforming declarative knowledge into procedural knowledge, which eventually leads to better detection accuracy. Our findings provide strong support for our major claim about the relative efficacy of feedback with concrete tips over feedback with abstract guidelines—at least in typical cases in which phishing cues are associated with low-level, technical attributes (e.g., fake links). However, mindful feedback makes individuals more conservative to make decision about an email's legitimacy. Decision avoidance has the potential to be beneficial when individuals encounter more advanced phishing strategies which they can't combat using their technical tips at hand. One of the significant contributions of this study to IS research lies in its highlighting of the importance of assimilating declarative forms of organizational guidelines into actionable forms of procedural sequences in the context of antiphishing training. Our theoretical framework is expected to offer valuable insights into an in-depth analysis of novel antiphishing programs.

Third, although feasibility is an important concept for understanding how people accomplish a given task (Bagozzi et al., 2003a, 2003b), it has rarely been incorporated into a model of antiphishing behavior and performance. Drawing on theories of feasibility, we are among the first to highlight the important role of perceived detection feasibility in antiphishing behavior and performance. In particular, our findings indicate that perceived detection feasibility remains significant in determining decision avoidance and detection accuracy even after controlling for the effects of coping adaptiveness (i.e., task-focused, emotion-focused, and avoidance coping) and other well-known factors. These results suggest that perceived feasibility is an essential factor for a better understanding of antiphishing behavior and performance.

Fourth, another contribution of this study to IS research lies in integrating the notion of feasibility with the concept of skill acquisition. Specifically, this paper provides a theoretical account on how feedback moderates the effect of perceived detection feasibility on antiphishing performance. We argue that because perceived detection feasibility is only a subjective behavioral possibility, proper antiphishing training could reinforce how such a possibility would successfully translate into antiphishing performance. More specifically, as compared with mindful feedback, technical feedback helps enable antiphishing tips to become procedural knowledge and thus successfully complete an antiphishing task. As theorized in our model, we found that the impact of perceived detection feasibility on detection accuracy is stronger after showing technical feedback than after mindful. This study is meaningful because it theoretically and empirically shows the important connection between the perspectives of feasibility and skill acquisition in the context of antiphishing training.

Finally, although much research exists concerning feedback characteristics in phishing research, little was known about whether feedback quantity makes a difference in antiphishing behavior and performance. Drawing on skill acquisition theory, this study is the first to shed light on the moderating effect of feedback quantity on the relationship between phishing cue saliency and its behavioral outcomes. Specifically, we theorize that for those with minimal technical information (i.e., low quantity feedback), phishing cue saliency is relatively less important as a determinant of detection accuracy. In contrast, for those with ample technical information (i.e., high quantity feedback), phishing cue saliency is more important as a factor regulating detection accuracy. Our findings clearly provide empirical support for the proposed model. In sum, this study is meaningful because it theoretically and empirically shows that the impacts of phishing characteristics on antiphishing outcomes vary considerably with how antiphishing feedback is provided.

5.3. Practical Contributions

Our study provides significant practical implications for information security practitioners who want to use feedback as part of their antiphishing training programs. First, our results emphasize the importance of technical feedback in improving detection accuracy. Specifically, technical feedback is shown to be crucial for helping people better identify the legitimacy of an email because it facilitates the formation of procedural knowledge related to phishing cues and antiphishing techniques. Thus, we encourage information security managers to consider incorporating technical feedback into their antiphishing training programs. However, mindful feedback helps individuals be conservative to make decision and it has the potential to prevent individuals from acting on a phishing email with more advanced techniques which they can’t combat with current technical tips. It can be used as supplementary education since it can facilitate antiphishing behavior in a different way compared to technical feedback. Second, as Figure 11 shows, more feedback does not seem always better in the context of antiphishing training. In particular, low quantity feedback is better for sophisticated phishing (low cue saliency), whereas, for simple phishing (high cue saliency), high quantity feedback is better. These findings do not necessarily advocate a specific amount of feedback as an ideal case. Instead, our study emphasizes the importance of ensuring that trainees thoroughly digest basic antiphishing tips before being exposed to other more sophisticated antiphishing techniques. In this manner, basic tips can be more efficiently translated into procedural knowledge in one’s mental representation, a mental step that can free up cognitive resources for assimilating new antiphishing techniques. In any case, because there is no magical one-shot training effective for all phishing messages, practitioners should have a clear idea on the pros and cons of their educational programs.

5.4. Limitations and Future Research Directions

Several limitations in our study need to be pointed out. First, all our experiments were based on web-based survey quizzes. Previous phishing studies have used similar online experiments to measure phishing detection accuracy (Sheng et al., 2010; Wang et al., 2016, 2017). This method has been recognized as an efficient approach for collecting data because it addresses ethical dilemmas that could result from field experiments (Wang et al., 2016).[12] Nevertheless, it would be helpful to reexamine the effects found in the present study in other nonexperimental settings. Second, we only tested the effects of feedback in a short term. The long-term effects of the feedback in the study remain to be tested. Third, the current study was focused only on phishing emails with deceitful links. Yet, other types of phishing attempts go far beyond such fraudulent links and are difficult to detect. For example, some phishing emails ask receivers to call certain phone numbers or reply to messages to further lure the receivers into providing information. Thus, our findings may not be applicable to other types of phishing. Fourth, our feedback material is relatively short, just one page. Thus, our findings regarding the role of feedback cannot be generalized to other longer forms of feedback. Lastly, we controlled for several factors when testing our hypotheses. However, we may still have failed to control some potentially important variables. Thus, our results should be interpreted carefully until more control variables are added into our model.

This study yields insights into several additional avenues for future research. First, it provides either technical or mindful feedback, not both, to each experimental group. Future research may test if providing both technical and mindful feedback will be more effective than using each of them separately. Second, although this study focused on phishing cue saliency as a phishing characteristic, future research should examine other phishing characteristics to better understand the actual ramifications that feedback has on antiphishing behavior and performance. For example, Goel et al. (2017) show that phishing attempts can be categorized by various factors such as the gain-loss frame, the extrinsic-intrinsic taxonomy, and personalization. We cannot expect that the effectiveness of a certain type of feedback is identical across these different types of phishing. Thus, we encourage researchers to further investigate the correspondence between feedback and phishing characteristics for more efficient and effective antiphishing training.

---

[12] Also, the phishing quizzes used in our study are ethically and practically useful for phishing awareness program. Please see https://www.washingtonpost.com/media/2020/09/23/tribune-bonus-email-phishing-hoax/.

5.5. Concluding Remarks

Despite the importance of offering proper feedback on individuals' reactions to phishing scams, our understanding is severely limited concerning the efficacy of such educational information under various types of phishing attacks. We developed and empirically tested a theoretical conceptual framework that describes how feedback and phishing characteristics interact to influence decision avoidance and detection accuracy. More research needs to be performed to untangle the complex interactions between feedback and phishing characteristics on antiphishing behavior and performance. We hope the proposed model will be helpful for such endeavors in this important line of research.

References

Al-Daeef MM, Basir N, Saudi MM (2017) Security awareness training: A review. Ao SI, Gelman L, Hukins DW, Hunter A, Korsunsky AM, eds. *Proc. World Congr. Eng.* 2017 (Newswood Limited, Hong Kong), 1(July):5-7.

Anderson CJ (2003) The psychology of doing nothing: Forms of decision avoidance result from reason and emotion. *Psych. Bull.* 129(1):139-167.

Anderson JR (1976) *Language, Memory, and Thought*, 1st ed. (Laurence Erlbaum Associates, Inc., New York).

Anderson JR (1982) Acquisition of cognitive skill. *Psych. Rev.* 89(4):369-406.

Anderson JR (1987) Skill acquisition: Compilation of weak-method problem situations. *Psych. Rev.* 94(2):192-210.

Anderson JR (2010) *Cognitive Psychology and Its Implications*, 7th ed. (Macmillan, New York).

Arachchilage NA, Love S, Beznosov K (2016) Phishing threat avoidance behavior: An empirical investigation. *Comput. Human Behav.* 60(July):185-197.

Bagozzi RP, Bergami M, Leone L (2003a) Hierarchical representation of motives in goal setting. *J. Appl. Psych.* 88(5):915-943.

Bagozzi RP, Dholakia UM, Basuroy S (2003b) How effortful decisions get enacted: The motivating role of decision processes, desires, and anticipated emotions. *J. Behav. Decision Mak*ing 16(4):273-295.

Bagozzi RP, Yi Y (1988) On the evaluation of structural equation models. *J. Acad. Marketing Sci.* 16(1):74-94.

Bangert-Drowns RL, Kulik CL, Kulik JA, Morgan M (1991) The instructional effect of feedback in test-like events. *Rev. Ed. Res.* 61(2):213-238.

Baron RM, Kenny DA (1986) The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *J. Personality Soc. Psych.* 51(6):1173-1182.

Bialystok E (1979) Explicit and implicit judgements of L2 grammaticality. *Lang. Learn*. 29(1):81-103.

Bilodeau IM (1966) Information feedback. Bilodeau EA, eds. *Acquisition of Skill* (Academic Press, New York), 255-296.

Braarud PØ (2001) Subjective task complexity and subjective workload: Criterion validity for complex team tasks. *Internat. J. Cogn. Ergonomics*. 5(3):261-273.

Butler R (1987) Task-involving and ego-involving properties of evaluation: Effects of different feedback conditions on motivational perceptions, interest, and performance. *J. Ed. Psych*. 79(4):474-482.

Butler R, Nisan M (1986) Effects of no feedback, task-related comments, and grades on intrinsic motivation and performance. *J. Ed. Psych*. 78(3):210-216.

Canova G, Volkamer M, Bergmann C, Borza R, Reinheimer B, Stockhardt S, Tenberg R (2015b) Learn to spot phishing URLs with the android NoPhish app. Bishop M, Miloslavskaya N, Theocharidou M, eds. *IFIP World Conf. Inform. Security Ed*. 2015 (Springer International Publishing, Cham), 87-100.

Carver CS, Scheier MF (1982) Control theory: A useful conceptual framework for personality-social, clinical, and health psychology. *Psych. Bull*. 92(1):111-135.

Cook DM (1968) The impact on managers of frequency of feedback. *Acad. Management J*. 11(3):263-277.

Devine DJ, Kozlowski SW (1995) Domain-specific knowledge and task characteristics in decision making. *Organ. Behav. Human Decision Processes*. 64(3):294-306.

Downs JS, Holbrook MB, Cranor LF (2006) Decision strategies and susceptibility to phishing. Cranor LF, ed. *Proc. 2nd Sympos. Usable Privacy Security. 2006* (Association for Computing Machinery, New York), 79-90.

Dutton JE (1988) Patterns of interest around issues: The role of uncertainty and feasibility. *Acad. Management J.* 31(3):663-675.

Earley PC, Northcraft GB, Lee C, Lituchy TR (1990) Impact of process and outcome feedback on the relation of goal setting to task performance. *Acad. Management J.* 33(1):87-105.

Educause (2020) Cybersecurity awareness resource library. Accessed June 24, 2020, https://www.educause.edu /focus-areas-and-initiatives/policy-and-security/cybersecurity-program/resources/information-security-guide/toolkits/cybersecurity-awareness-resource-library.

Federal Bureau of Investigation (2019) Internet crime report. Report, Internet Crime Complaint Center, Federal Bureau of Investigation, Washington, DC.

Fitzsimmons JR, Douglas EJ (2011) Interaction between feasibility and desirability in the formation of entrepreneurial intentions. *J. Bus. Ventur*. 26(4):431-440.

Fornell C, Larcker DF (1981) Evaluating structural equation models with unobservable variables and measurement error. *J. Marketing Res*. 18(1):39-50.

Gefen D, Straub DW (2000) The relative importance of perceived ease of use in IS adoption: A study of e-commerce adoption. *J. Assoc. Inform. Systems*. 1(1):8.

Goel S, Williams K, Dincelli E (2017) Got phished? Internet security and human vulnerability. *J. Assoc. Inform. Systems.* 18(1):22-44.

Griffin RW, Welsh A, Moorhead G (1981) Perceived task characteristics and employee performance: A literature review. *Acad. Management Rev*. 6(4):655-664.

Harks B, Rakoczy K, Hattie J, Besser M, Klieme E (2014) The effects of feedback on achievement, interest and self-evaluation: The role of feedback's perceived usefulness. *Ed. Psych*. 34(3):269-290.

Hattie J, Timperley H (2007) The power of feedback. *Rev. Ed. Res.* 77(1):81-112.

Heckhausen H (2013) *The Anatomy of Achievement Motivation* (Academic Press, New York and London).

Jampen D, Gür G, Sutter T, Tellenbach B (2020) Don't click: towards an effective anti-phishing training. A comparative literature review. *Human-centric Comput. Inform. Sci.* 10(1):1-41.

Jensen ML, Dinger M, Wright RT, Thatcher JB (2017) Training to mitigate phishing attacks using mindfulness techniques. *J. Management Inform. Systems*. 34(2):597-626.

Jung JH, Schneider C, Valacich J (2010) Enhancing the motivational affordance of information systems: The effects of real-time performance feedback and goal setting in group collaboration environments. *Management Sci.* 56(4):724-742.

Kahneman D (1973) *Attention and Effort* (Prentice-Hall inc., Englewood Cliffs, NJ).

Kanfer R, Ackerman PL (1989) Motivation and cognitive abilities: An integrative/aptitude-treatment interaction approach to skill acquisition. *J. Appl. Psych*. 74(4):657-690.

Kim S, Stoel L (2004) Apparel retailers: Website quality dimensions and satisfaction. *J. Retailing Consumer Services*. 11(2):109-117.

Klein JI (1990) Feasibility theory: A resource-munificence model of work motivation and behavior. *Acad. Management Rev.* 15(4):646-665.

Kluger AN, DeNisi A (1996) The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psych. Bull*. 119(2):254-284.

Kluger AN, Lewinsohn S, Aiello JR (1994) The influence of feedback on mood: Linear effects on pleasantness and curvilinear effects on arousal. *Organ. Behav. Human Decision Processes*. 60(2):276-299.

Komaki J, Heinzmann AT, Lawson L (1980) Effect of training and feedback: Component analysis of a behavioral safety program. *J. Appl. Psych*. 65(3):261-270.

Kumaraguru P, Rhee Y, Acquisti A, Cranor LF, Hong J, Nunge E (2007a) Protecting people from phishing: The design and evaluation of an embedded training email system. Rosson MB, ed. *Proc. SIGCHI Conf. Human Factors Comput. Systems. 2007* (Association for Computing Machinery, New York), 905-914.

Lam CF, DeRue DS, Karam EP, Hollenbeck JR (2011) The impact of feedback frequency on learning and task performance: Challenging the "more is better" assumption. *Organ. Behav. Human Decision Processes*. 116(2):217-228.

Liberman N, Trope Y (1998) The role of feasibility and desirability considerations in near and distant future decisions: A test of temporal construal theory. *J. Personality Soc. Psych*. 75(1):5-18.

Liu P, Li Z (2012) Task complexity: A review and conceptualization framework. *Internat. J. Indust. Ergonomics*. 42(6):553-568.

Mohammed S, Harrison DA (2013) The clocks that time us are not the same: A theory of temporal diversity, task characteristics, and performance in teams. *Organ. Behav. Human Decision Processes*. 122(2):244-256.

Newell KM (1976) Knowledge of results and motor learning. *Exercise Sport Sci. Rev*. 4(1):195–228.

Norman DA, Bobrow DG (1975) On data-limited and resource-limited processes. *Cogn. Psych*. 7(1):44-64.

Oldham GR, Hackman JR, Pearce JL (1976) Conditions under which employees respond positively to enriched work. *J. Appl. Psych*. 61(4):395-403.

Osterman Research (2019) The ROI of security awareness training. Report, Osterman Research, Black Diamond, WA.

Palmer JW (2002) Web site usability, design, and performance metrics. *Inform. Systems Res*. 13(2):151-167.

PhishLabs (2019) Phishing trends and intelligence report. Report, PhishLabs, Charleston, SC.

Proofpoint (2020) State of the phish report. Report, Proofpoint, Sunnyvale, CA.

Rakoczy K, Klieme E, Bürgermeister A, Harks B (2008) The interplay between student evaluation and instruction: Grading and feedback in mathematics classrooms. *J. Psych.* 216(2):111-124.

Robinson P (2001) Task complexity, task difficulty, and task production: Exploring interactions in a componential framework. *Appl. linguist*. 22(1):27-57.

Roch SG, Lane JA, Samuelson CD, Allison ST, Dent JL (2000) Cognitive load and the equality heuristic: A two-stage model of resource overconsumption in small groups. *Organ. Behav. Human Decision Processes*. 83(2):185-212.

Salmoni AW, Schmidt RA, Walter CB (1984) Knowledge of results and motor learning: A review and critical reappraisal. *Psych. Bull*. 95(3):355-386.

Schuetz SW, Benjamin Lowry P, Pienta, DA, Bennett Thatcher J (2020) The effectiveness of abstract versus concrete fear appeals in information security. *J. Management Inform. Systems*. 37(3):723-757.

Segars AH, Grover V (1998) Strategic information systems planning success: An investigation of the construct and its measurement. *MIS Quart*. 22(2):139-163.

Sheng S, Holbrook M, Kumaraguru P, Cranor LF, Downs J (2010) Who falls for phish? A demographic analysis of phishing susceptibility and effectiveness of interventions. Elizabeth DM, ed. *Proc. SIGCHI Conf. Human Factors Comput. Systems.* 2010 (Association for Computing Machinery, New York), 373-382.

Sheng S, Magnien B, Kumaraguru P, Acquisti A, Cranor LF, Hong J, Nunge E (2007) Anti-phishing phil: The design and evaluation of a game that teaches people not to fall for phish. Cranor LF, ed. *Proc. 3rd Sympos. Usable Privacy Security.* 2007 (Association for Computing Machinery, New York), 88-99.

Shepherd LA, Archibald J (2017) Security awareness and affective feedback: Categorical behavior vs. reported behavior. Onwubiko C, ed. *Internat. Conf. Cyber Situational Awareness, Data Anal. Assessment.* 2017 (Institute of Electrical and Electronics Engineers, Washington, DC), 1-6.

Silic M, Lowry PB (2020) Using design-science based gamification to improve organizational security training and compliance. *J. Management Inform. Systems*. 37(1):129-161.

Steele-Johnson D, Steinke J, Kalinoski Z (2011) Cognitive ability and objective and subjective task complexity: Unique and differential effects on performance, self-efficacy, and cognitive appraisals. *J. Organ. Psych*. 11(1):73-86.

Stockhardt S, Reinheimer B, Volkamer M, Mayer P, Kunz A, Rack P, Lehmann D (2016) Teaching phishing-security: Which way is best? Hoepman JH, Katzenbeisser S, eds. *IFIP Internat. Conf. ICT Systems Security Privacy Protection.* 2016 (Springer International Publishing, Cham), 135-149.

Taatgen NA, Van Rijn H, Anderson J (2007) An Integrated Theory of Prospective Time Interval Estimation: The Role of Cognition, Attention, and Learning. *Psych. Rev.* 114(3):577-598.

Tseng ST, Levy PE, Young SHA, Thibodeau RK, Zhang, X (2019) Frequent feedback in modern organizations: Panacea or fad? Steelman LA, Williams JR, eds. *Feedback at Work* (Springer International Publishing, Cham), 53-73.

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

Verizon (2020) 2020 Data Breach Investigations Report. Report, Verizon, New York.

Vroom VH (1964) *Work and Motivation* (Wiley, New York).

Wang J, Li Y, Rao HR (2016) Overconfidence in Phishing Email Detection. *J. Assoc. Inform. Systems* 17(11):759 -783.

Wang J, Li Y, Rao HR (2017) Coping responses in phishing detection: An investigation of antecedents and consequences. *Inform. Systems Res*. 28(2):378-396.

Waterloo News (2019) Why people don't reply to your emails? *Waterloo News* (April 3), https://uwaterloo.ca/news/news/why-people-dont-reply-your-emails.

Wen ZA, Lin Z, Chen R, Andersen E (2019, May) What.hack: Engaging anti-phishing training through a role-playing phishing simulation game. Brewster SA, Fitzpatrick G, eds. *Proc. 2019 CHI Conf. on Human Factors Comput. Systems.* 2019 (Association for Computing Machinery, New York), 1-12.

Werbach K, Hunter D (2012) *For the Win: How Game Thinking Can Revolutionize Your Business* (Wharton Digital Press, Philadelphia, PA).

Wickens CD (1984) Processing resources in attention. Parasuraman R, Davies DR, eds. *Varieties of Attention* (Academic Press, New York), 63-102.

Wright RT, Jensen ML, Thatcher JB, Dinger M, Marett K (2014) Research note-Influence techniques in phishing attacks: An examination of vulnerability and resistance. *Inform. Systems Res*. 25(2):385-400.

Xu J, Benbasat I, Cenfetelli RT (2014) Research note-The influences of online service technologies and task complexity on efficiency and personalization. *Inform. Systems Res*. 25(2):420-436.

Appendix C: Measurement Items

Unless specified, all items were measured with a 7-point Likert scale (1 - Strongly Disagree, 7 - Strongly Agree)

Manipulation Checks

Technical vs. Mindful Feedback
For detecting phishing emails, the previous education provides technical information.

Feedback Quantity
How many tips for detecting phishing emails have you seen in the previous education?
(One tip/Many tips)
One tip = 0, Many tips = 1

Research Variables

Age
Age in years

Gender (GEN)
Male = 0, Female = 1

Prior Phishing Experience (PPE)
How many times have you been phished in the past?
(5-point scales anchored with "None at all" and "A great deal")

Average Email Load (AEL)
Number of emails received per day on average

Preliminary Training Detection Accuracy (PTDA)
Is this a Phishing email? *(Yes/No/Skip)*
Detection Accuracy: Incorrect = -1, Skip = 0, Correct = 1

Task-Focused Coping (TC)
TC1. I made every effort to perform my goals.
TC2. I concentrated hard on doing well.

Emotion-Focused Coping (EC)
EC1. I worried about my inadequacies.
EC2. I blamed myself for not doing better.
EC2. I blamed myself for not knowing what to do.

Avoidance Coping (AC)
AC1. I acted as though the task wasn't important.
AC2. I didn't take the task too seriously.
AC2. I decided there was no point in trying to do well.

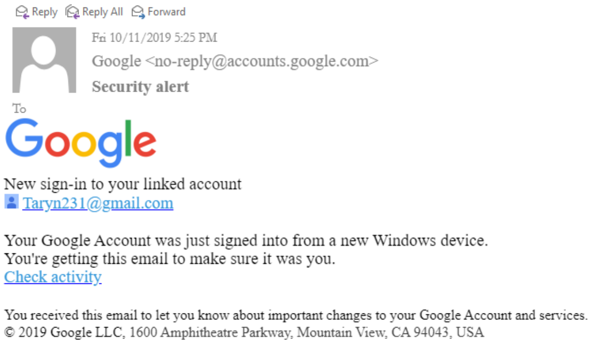
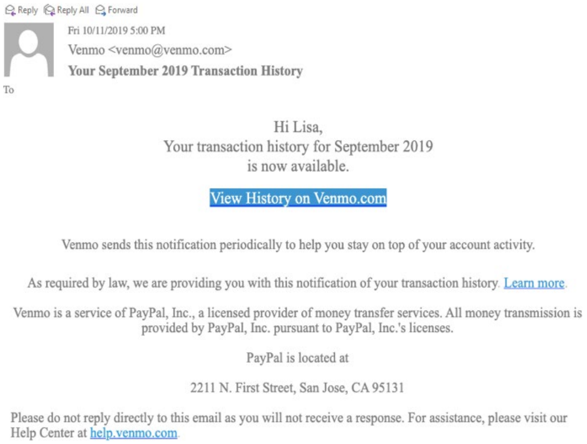
Perceived Detection Feasibility (PDF)
PDF1. It is possible to determine whether the email is phishing.
PDF2. It is feasible to determine whether the email is phishing.

PDF3. I am certain about my judgement of this email.
PDF4. I am sure of my judgement of this email.

Antiphishing Outcomes (DAV and DAC)
Is this a Phishing email? *(Yes/No/Skip)*
Decision Avoidance: Decision Making = 0, Skip = 1
Detection Accuracy: Incorrect = -1, Skip = 0, Correct = 1

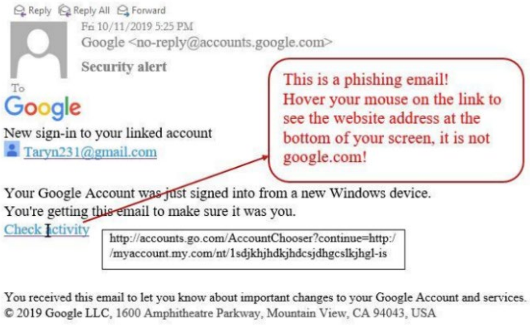
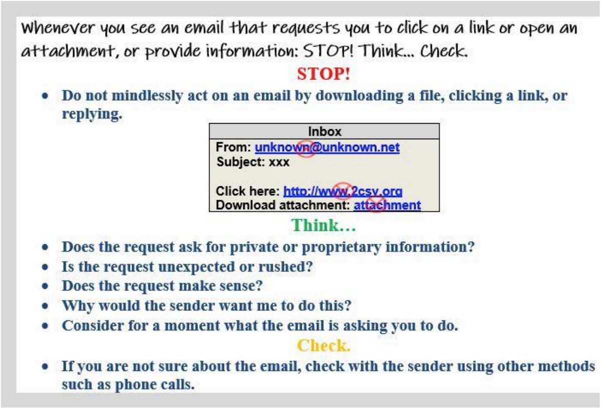
Appendix D: Email Samples and Manipulations

Email Samples[13]

Phishing Email	Legitimate Email
 Fri 10/11/2019 5:25 PM Google <no-reply@accounts.google.com> **Security alert** To New sign-in to your linked account Taryn231@gmail.com Your Google Account was just signed into from a new Windows device. You're getting this email to make sure it was you. Check activity You received this email to let you know about important changes to your Google Account and services. © 2019 Google LLC, 1600 Amphitheatre Parkway, Mountain View, CA 94043, USA	 Fri 10/11/2019 5:00 PM Venmo <venmo@venmo.com> **Your September 2019 Transaction History** To Hi Lisa, Your transaction history for September 2019 is now available. Venmo sends this notification periodically to help you stay on top of your account activity. As required by law, we are providing you with this notification of your transaction history. Learn more. Venmo is a service of PayPal, Inc., a licensed provider of money transfer services. All money transmission is provided by PayPal, Inc. pursuant to PayPal, Inc.'s licenses. PayPal is located at 2211 N. First Street, San Jose, CA 95131 Please do not reply directly to this email as you will not receive a response. For assistance, please visit our Help Center at help.venmo.com.

Manipulations in Experiment 1 and 2

[Feedback Type]

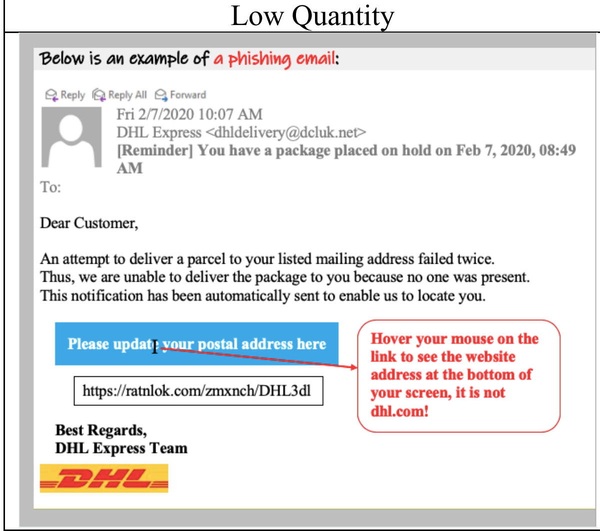
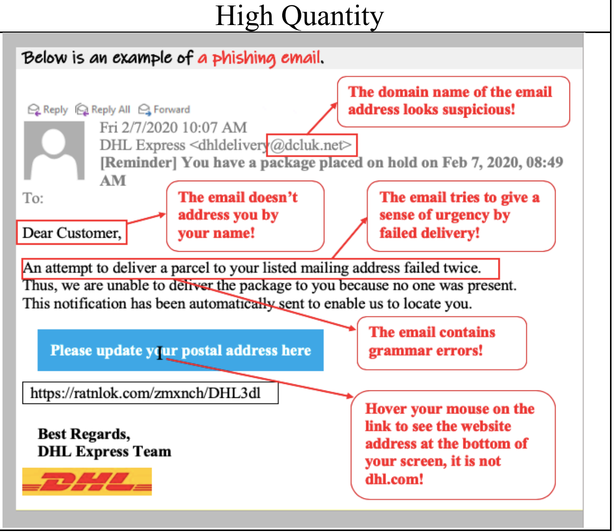
Technical Feedback	Mindful Feedback
 Fri 10/11/2019 5:25 PM Google <no-reply@accounts.google.com> **Security alert** To New sign-in to your linked account Taryn231@gmail.com Your Google Account was just signed into from a new Windows device. You're getting this email to make sure it was you. Check activity http://accounts.go.com/AccountChooser?continue=http://myaccount.my.com/nt/1sdjkhjhdkjhdcsjdhgcslkjhgl-is This is a phishing email! Hover your mouse on the link to see the website address at the bottom of your screen, it is not google.com! You received this email to let you know about important changes to your Google Account and services. © 2019 Google LLC, 1600 Amphitheatre Parkway, Mountain View, CA 94043, USA	 Whenever you see an email that requests you to click on a link or open an attachment, or provide information: STOP! Think… Check. **STOP!** • **Do not mindlessly act on an email by downloading a file, clicking a link, or replying.** From: unknown@unknown.net Subject: xxx Click here: http://www.2csv.org Download attachment: attachment **Think…** • **Does the request ask for private or proprietary information?** • **Is the request unexpected or rushed?** • **Does the request make sense?** • **Why would the sender want me to do this?** • **Consider for a moment what the email is asking you to do.** **Check.** • **If you are not sure about the email, check with the sender using other methods such as phone calls.**

---

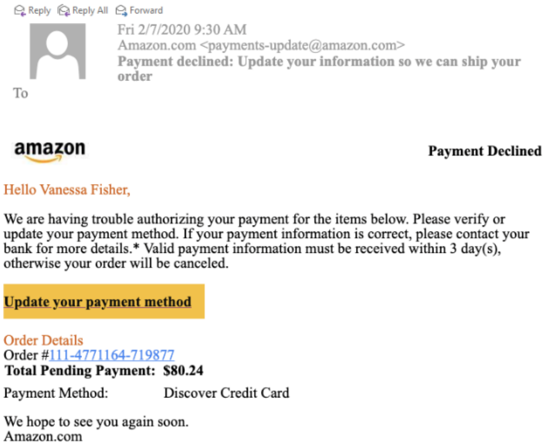
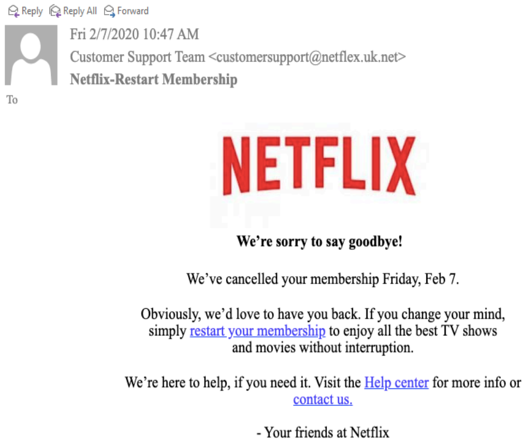
[13] In Experiment 3, we manipulated phishing cue saliency of the emails. Please refer to "Manipulations in Experiment 3" for detailed information.

Manipulations in Experiment 3

[Feedback Quantity]

Low Quantity	High Quantity
 Below is an example of a phishing email: Fri 2/7/2020 10:07 AM DHL Express <dhldelivery@dcluk.net> **[Reminder] You have a package placed on hold on Feb 7, 2020, 08:49 AM** To: Dear Customer, An attempt to deliver a parcel to your listed mailing address failed twice. Thus, we are unable to deliver the package to you because no one was present. This notification has been automatically sent to enable us to locate you. **Please update your postal address here** https://ratnlok.com/zmxnch/DHL3dl **Best Regards,** **DHL Express Team** **Hover your mouse on the link to see the website address at the bottom of your screen, it is not dhl.com!**	 Below is an example of a phishing email. Fri 2/7/2020 10:07 AM DHL Express <dhldelivery@dcluk.net> **[Reminder] You have a package placed on hold on Feb 7, 2020, 08:49 AM** To: Dear Customer, An attempt to deliver a parcel to your listed mailing address failed twice. Thus, we are unable to deliver the package to you because no one was present. This notification has been automatically sent to enable us to locate you. **Please update your postal address here** https://ratnlok.com/zmxnch/DHL3dl **Best Regards,** **DHL Express Team** **The domain name of the email address looks suspicious!** **The email doesn't address you by your name!** **The email tries to give a sense of urgency by failed delivery!** **The email contains grammar errors!** **Hover your mouse on the link to see the website address at the bottom of your screen, it is not dhl.com!**

[Phishing Cue Saliency][14]

Phishing with Low Cue Saliency	Phishing with High Cue Saliency
 Fri 2/7/2020 9:30 AM Amazon.com <payments-update@amazon.com> **Payment declined: Update your information so we can ship your order** To **Payment Declined** Hello Vanessa Fisher, We are having trouble authorizing your payment for the items below. Please verify or update your payment method. If your payment information is correct, please contact your bank for more details.* Valid payment information must be received within 3 day(s), otherwise your order will be canceled. **Update your payment method** Order Details Order #111-4771164-719877 **Total Pending Payment: $80.24** Payment Method: Discover Credit Card We hope to see you again soon. Amazon.com	 Fri 2/7/2020 10:47 AM Customer Support Team <customersupport@netflex.uk.net> **Netflix-Restart Membership** To **We're sorry to say goodbye!** We've cancelled your membership Friday, Feb 7. Obviously, we'd love to have you back. If you change your mind, simply restart your membership to enjoy all the best TV shows and movies without interruption. We're here to help, if you need it. Visit the Help center for more info or contact us. - Your friends at Netflix

---

[14] For legitimate emails, we couldn't manipulate the appearance of an email. Thus, we chose legitimate emails with very short URLs such as "https://venmo.com" as in the high phishing cue saliency group and chose legitimate emails with relatively long URLs such as "https://buy.itunes.apple.com" as in the low phishing cue saliency group.

Appendix E: Descriptive Statistics, Composite Reliabilities, AVEs, and Correlation Matrix

Table 14. Experiment 2

	Mean	SD	CR	AVE	Correlation Matrix												
					1	2	3	4	5	6	7	8	9	10	11	12	13
Individual Level																	
1 FT	0.50	0.50	-	-	1.00												
2 Age	46.44	13.27	-	-	0.05	1.00											
3 GEN	0.47	0.50	-	-	0.04	-0.33	1.00										
4 PPE	2.69	1.28	-	-	-0.01	0.20	-0.14	1.00									
5 AEL	47.44	43.62	-	-	0.00	0.21	0.14	0.04	1.00								
6 PTDA	-0.02	0.72	-	-	-0.09	0.29	-0.03	0.24	0.10	1.00							
7 TC	6.44	0.83	0.87	0.77	0.08	-0.05	0.01	-0.09	0.09	0.04	1.00						
8 EC	3.86	1.70	0.87	0.70	-0.10	-0.22	0.07	-0.05	0.00	-0.37	0.05	1.00					
9 AC	1.93	1.30	0.92	0.80	-0.10	0.00	0.06	0.22	-0.02	0.05	-0.40	0.17	1.00				
Message level																	
10 PDF	5.44	1.19	0.94	0.80	0.05	0.03	-0.13	0.13	0.11	0.15	0.31	-0.13	0.01	1.00			
11 RT	28.44	27.30	-	-	-0.05	0.14	-0.08	0.03	-0.08	0.16	0.07	-0.09	-0.20	-0.11	1.00		
12 DAV	0.02	0.14	-	-	-0.11	-0.04	0.12	0.06	0.14	-0.01	0.01	0.15	0.20	-0.12	-0.02	1.00	
13 DAC	0.33	0.93	-	-	0.18	0.01	-0.07	-0.08	-0.02	-0.05	0.04	-0.04	-0.09	0.11	0.01	-0.05	1.00

Notes:

- SD = standard deviation; CR = composite reliability; AVE = average variance extracted.
- FT = feedback type (mindful = 0, technical = 1); GEN = gender (male = 0, female = 1); PPE = prior phishing experience; AEL = average email load; PTDA = preliminary training detection accuracy; TC = task-focused coping; EC = emotion-focused coping; AC = avoidance coping; PDF = perceived detection feasibility; RT = response time to each quiz; DAV = decision avoidance; DAC = detection accuracy.

Table 15. Experiment 3

	Mean	SD	CR	AVE	Correlation Matrix													
					1	2	3	4	5	6	7	8	9	10	11	12	13	14
Individual Level																		
1 FQ	0.56	0.50	-	-	1.00													
2 Age	38.36	9.39	-	-	-0.03	1.00												
3 GEN	0.51	0.50	-	-	0.02	0.13	1.00											
4 PPE	2.20	0.99	-	-	0.03	0.21	0.20	1.00										
5 AEL	39.17	62.85	-	-	0.08	0.05	0.03	0.12	1.00									
6 PTDA	0.16	0.74	-	-	0.07	0.08	-0.12	0.01	0.04	1.00								
7 TC	6.71	0.52	0.72	0.57	-0.06	0.12	0.10	0.10	0.05	-0.01	1.00							
8 EC	2.92	1.70	0.89	0.74	-0.02	0.13	0.26	0.14	-0.02	-0.22	0.00	1.00						
9 AC	1.28	0.67	0.95	0.85	0.01	-0.12	-0.14	-0.03	-0.07	-0.05	-0.51	0.10	1.00					
Message Level																		
10 PCS	0.50	0.50	-	-	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00				
11 PDF	5.62	1.37	0.95	0.83	-0.03	-0.05	-0.09	-0.04	0.02	0.17	0.14	-0.27	-0.12	0.08	1.00			
12 RT	21.80	19.35	-	-	0.08	0.07	0.04	0.05	0.01	0.03	0.07	0.02	-0.04	-0.07	-0.13	1.00		
13 DAV	0.02	0.14	-	-	0.00	-0.01	-0.02	0.00	-0.02	-0.02	-0.03	0.01	0.01	0.00	-0.27	-0.01	1.00	
14 DAC	0.57	0.81	-	-	-0.06	0.03	-0.06	-0.05	0.01	0.02	0.00	-0.14	-0.06	0.15	0.24	-0.05	-0.10	1.00

Notes:

- SD = standard deviation; CR = composite reliability; AVE = average variance extracted.
- FQ = feedback quantity (low = 0, high = 1); GEN = gender (male = 0, female = 1); PPE = prior phishing experience; AEL = average email load; PTDA = preliminary training detection accuracy; TC = task-focused coping; EC = emotion-focused coping; AC = avoidance coping; PCS = phishing cue saliency (low = 0, high = 1); PDF = perceived detection feasibility; RT = response time to each quiz; DAV = decision avoidance; DAC = detection accuracy.