The Roles of Knowledge, Cue Utilisation, and Decision Styles in Phishing Email Detection

This thesis is submitted in partial fulfilment of the Honours degree of Bachelor of Psychological Science (Honours)

School of Psychology

The University of Adelaide

September 2022

Word Count: 9,480

Table of Contents

List of Figures

List of Tables

Abstract

Phishing emails are one of the most pervasive and costly threats to cybersecurity worldwide. To mitigate this risk, email users need to be able to accurately detect these fraudulent emails. Although many knowledge-based training programs improve detection, a considerable proportion of users still fail to identify phishing emails after training. As such, there needs to be a greater understanding of the factors which may enhance discrimination between genuine and phishing emails. While the role of cognitive factors has previously been examined, few studies have made the important distinction between *detection* and *discrimination* when investigating knowledge, cue utilisation and decision styles. An age-stratified sample of Australian residents ($N = 144$) completed an online phishing email detection task, and measures of objective phishing email knowledge, cue utilisation, and intuitive and rational decision styles. While both higher knowledge and cue utilisation were associated with greater *detection* of phishing emails, only cue utilisation was associated with a greater capacity for *discrimination* between genuine and phishing emails. Knowledge was instead associated with greater caution. An intuitive style was associated with poorer detection of phishing emails, with no relationship found between a rational style and detection performance. Overall, these findings suggest that increasing knowledge may not be sufficient to provide users with the ability to discriminate between genuine and phishing emails. Increasing cue utilisation, in addition to knowledge, may be a more effective approach. Practically, the present study provides evidence that users within broader society may benefit from cue-based training to enhance phishing email detection.

Keywords: Phishing email detection, cue utilisation, knowledge, decision styles, discrimination

Declaration

"This thesis contains no material which has been accepted for the award of any other degree of diploma in any University, and, to the best of my knowledge, this thesis contains no material previously published except where due reference is made. I give permission for the digital version of this thesis to be made available on the web, via the University of Adelaide's digital thesis repository, the Library Search and through web search engines, unless permission has been granted by the School to restrict access for a period of time."

September 2022

Student Contribution to Experimental Design

The student and supervisors collaborated in the development of research questions and the student was supported by the supervisors in the selection of the variables which were investigated in the present study. The student and supervisors also collaborated in the conceptual design of the Email Sorting Task. The student was not involved in the development of the Objective Phishing Email Knowledge scale or the adapted Decision Styles Scale. The student was responsible for uploading all experimental materials into the Qualtrics survey platform. The data was collected by the supervisors via the Qualtrics platform. The student collated, cleaned, and prepared the data for analysis. While supported by the supervisors, all analyses were run by the student. The student was responsible for all written components of the thesis.

The Roles of Knowledge, Cue Utilisation, and Decision Styles in Phishing Email Detection

Phishing is a form of cyber-attack which uses deception, and often impersonation, to obtain access to a victim's sensitive information (Lastdrager, 2014). While phishing messages can be sent across many platforms, they are particularly common and pervasive through email (Chaudry et al., 2016; Jampen et al., 2020). Phishing emails typically attempt to masquerade as legitimate messages, mimicking the visual look of sources that would otherwise be familiar or non-threatening to the victim (Zhuo et al., 2022). These fraudulent emails often feature a malicious link which re-directs the victim to a fake landing page where they are prompted to provide their sensitive details (Jampen et al., 2020).

Phishing attacks are one of the greatest global cybersecurity threats for organisations, governments, and individuals (Salahdine & Kaabouch, 2019; Zhuo et al., 2022). In 2021, phishing attacks victimized approximately 90% of surveyed Australian organisations and was the most reported form of cyber attack in Australia (Proofpoint, 2022; Australian Competition and Consumer Comission, 2022). For organisations within the United States, the average annual costs associated with phishing emails has more than tripled since 2015, reaching $14.8 million in 2021 (Ponemon Institute; 2021). In addition to financial consequences, phishing attacks can also compromise the privacy of individuals and damage the psychological wellbeing and cyber confidence of phishing victims (Jansen & Leukfedlt, 2018).

Despite the widespread use technological interventions that aim to reduce the incidence of phishing attacks, such as email filters, human email users continue to be exposed to phishing emails (Chaudry et al., 2016; Dou et al., 2017). It has been widely demonstrated that users typically have an extremely limited capacity to detect phishing emails and discriminate them from genuine emails (Constantino et al., 2018; Parsons et al., 2019; Sarno

& Neider 2021) Thus, users are often considered the 'weak link' in cybersecurity (Althobaiti et al., 2021; Parsons et al., 2019) and yet, they are the last and most critical defence against phishing email attacks (Khonji et al., 2013; Parsons et al., 2017). Therefore, understanding the factors which may influence a user's capacity to detect phishing emails is necessary to inform the development and application of effective phishing training programs (Das et al., 2020; Zhuo et al., 2022).

There has been considerable research investigating the impact of various demographic, social, and psychological factors on phishing email detection (Alseadoon et al., 2014; Greitzer at al., 2021; Lin et al., 2019). While there are some inconsistent findings, collectively the evidence suggests there are many individual differences which may be influential, such as personality and age (Darwish et al., 2012; Sheng et al., 2010; Zhuo et al., 2022). Recently, literature has focused on investigating cognitive factors which could be influenced via training to improve detection performance, such as cue utilisation, decision styles, and knowledge (Parsons et al., 2019; Sturman et al., 2022; Zhuo et al., 2022).

In the phishing email context, knowledge has been widely considered the foundation for detection and has been conceptualised in a variety of ways (Wang et al., 2012; Yang et al., 2022; Zhuo et al., 2022). For instance, as knowledge of the predictive features of a phishing email (Harrison et al., 2016), of the existence or prevalence of phishing emails (Diaz et al., 2019), and of the definition of phishing (Vishwanath et al., 2011). Phishing email knowledge has been consistently shown to be associated with reduced phishing victimisation and enhanced detection (Kumaraguru et al., 2010; Vishwanath et al., 2011; Wang et al., 2012; Yang et al., 2022). Consequently, many phishing training programs aim to increase knowledge of phishing emails (Al-Daeef et al., 2017; Kumaraguru et al., 2010; Mayhorn & Neyeste, 2012; Sumner & Yuan, 2019).

Knowledge-based training programs are designed to enhance users' capacity to identify phishing emails and distinguish them from genuine emails (Al-Daeef et al., 2017; Jampen et al., 2020; Kumaraguru et al., 2010). However, there are mixed results regarding the effectiveness of such training (Caputo et al., 2013; Harrison, 2018; Jampen et al., 2020). For example, Sheng et al. (2010) found that although training reduced phishing email victimisation by 40%, participants still failed to detect phishing emails 28% of the time. Further, a recent report found 65% of organisations which had been phished had previously trained their staff (Cloudian, 2021). Considering this evidence, it appears the available phishing email training programs may not be as effective as the literature suggests they could be (Harrison, 2018; Jampen et al., 2022; Zhuo et al., 2022).

There are a few possible explanations that could account for the relatively low success rate of such training programs. One factor may be that, within the literature, phishing email knowledge is defined and operationalised inconsistently. This may cause variation in the way the relationship between knowledge and detection performance is understood. Thus, it may be difficult to use empirical research to inform the construction of effective phishing training programs (Jampen et al., 2020).

Many studies define phishing email knowledge as an awareness of the prevalence or definition of phishing emails (Alnajim & Munro, 2009; Downs et al., 2007; Musuva et al., 2019). For example, Diaz et al. (2019) defined knowledge as having an awareness that phishing emails exist. However, such generic definitions of knowledge may be limited in their usefulness. For instance, an individual could be aware of the existence, prevalence, and definition of phishing, without understanding how to effectively discriminate between a genuine and phishing email.

Comparatively fewer studies have defined phishing email knowledge as an understanding of the specific features of an email which predict the likelihood that an email is

fraudulent (Harrison et al., 2016; Parsons et al., 2019). Features such as suspicious URL links and spelling and grammatical errors are frequently contained in phishing emails and can be used to determine if an email is legitimate (Althobaiti et al., 2021; Lötter & Futcher, 2015; Parsons et al., 2016). Some researchers have suggested the effective detection of phishing emails relies on the individuals' ability to identify these predictive features (Harrison et al., 2016; Walsh, 2020; Williams et al., 2018). Therefore, measuring knowledge of the predictive features of phishing emails may be beneficial when investigating detection performance.

Additionally, the way phishing email knowledge has been operationalised within phishing email detection literature has limitations. Many studies use measures which are subjective (Musuva et al., 2019; Sturman et al., 2022; Wang et al., 2016). For instance, Parsons et al. (2019) measures knowledge using the Human-Aspects-of-Information-Security-Questionnaire (HAIS-Q) which prompts self-reported agreement with a series of statements (e.g., "I don't open email attachments if the sender is unknown to me"; Parsons et al., 2017). A subjective approach prompts a self assessment which may be unlikely to be an accurate reflection of an individual's true level of phishing email knowledge (Canfield et al., 2019; Downs et al., 2007; Harrison et al., 2016).

A small number of studies use an objective operationalisation of phishing email knowledge, such as in the form of a test or performance assessment. Further, the few scales which have been developed may be limited in scope or reliability (Arachchilage & Love, 2014; Harrison et al., 2016). Overall, there is a distinct lack of empirical research which measures the relationship between objective phishing email knowledge and detection performance. As such, this limits the understanding of the role of knowledge in phishing email detection.

An additional limitation of much of the literature examining phishing email detection is that many studies only measure hits (e.g., when a phishing email is correctly identified as

phishing; Canfield & Fischhoff, 2018). Although this approach may be intuitive, it fails to explain whether improvements in detection performance are due to a greater overall bias towards classifying emails as phishing, or a greater capacity to discriminate between genuine and phishing emails (Butavicius et al., 2016; Canfield & Fischhoff, 2018).

Operationalising detection performance according to Signal Detection Theory (Stanislaw & Torodov, 1999) is a more robust methodology which measures the ability to discriminate between a stimulus of interest (e.g., phishing emails) and noise (e.g., genuine emails). By measuring false alarms (e.g., incorrect classifications of genuine emails as phishing) as well as hits, sensitivity (discrimination) and decision criterion (response bias) can be separately quantified. This distinction is important, as changes in either discrimination or response bias can similarly lead to a greater hit rate, despite having different underlying implications (Canfield et al., 2016).

For instance, with a response bias towards phishing classifications, an individual indiscriminately classifies an increased number of emails as phishing. This cautious approach to email use results in increased hits, but also increased false alarms (Canfield et al., 2016). Alternatively, possessing a greater capacity for discrimination between genuine and phishing emails may be more advantageous (Parsons et al., 2019), as detection improves without increasing false alarms. Thus, while greater detection can be caused by more cautious behaviour, increased discrimination reflects a greater ability to identify the differences between genuine and phishing emails. Therefore, when only detection (hits) is measured, it remains unclear whether greater detection performance is caused by increased caution or discrimination (Canfield & Fischhoff, 2018).

Although a distinction between detection and discrimination has been increasingly made within phishing email research (Canfield et al., 2016; Sarno et al., 2019; Sarno & Neider, 2021), the few studies which investigate the role of phishing email knowledge record

mixed results. While some demonstrate knowledge is associated with enhanced discrimination (Parsons et al., 2017; Sturman et al., 2022), others indicate possessing higher knowledge may only make email users more cautious (Anandpara et al., 2007; Kumaraguru et al., 2007; Sheng et al., 2010). Consequently, it remains unclear whether increasing knowledge improves detection by influencing discrimination or eliciting a response bias (Harrison, 2018; Sarno et al., 2022). Given the clear implications of these outcomes on the effectiveness of training, future research is warranted.

Another possible explanation for why some users continue to fail to detect phishing emails after training is that, although knowledge is necessary for high level performance in a skill, it is not sufficient (Kahneman & Klein, 2009; Rasmussen, 1986; Sturman et al., 2022). Therefore, increasing phishing email knowledge alone may not lead to improved discrimination. Knowledge-based training typically imparts explicit knowledge (e.g., strategies to be used to assess email legitimacy; Harrison, 2018) which individuals can consciously repeat and apply (Zhuo et al., 2022). However, the behaviours associated with this crystalised form of knowledge are characteristic of novice performance (Rasmussen, 1986).

Rasmussen's (1986) model of skill acquisition describes how individuals progress from novice to expert through three stages of performance: from knowledge-based, to rule-based, to skill-based behaviour. The first, knowledge-based stage involves slow, analytical, and effortful processing to make sense of novel patterns and stimuli (Rasmussen, 1986). A novice email user who tries to detect phishing emails will engage in this behaviour, where they may concentrate heavily on the task and consciously consider a broad range of email features to inform their assessment of an email's legitimacy. Rule-based behaviour is demonstrated when sufficient experience is gained to intentionally follow a mental template

of optimal performance (e.g., following a step-by-step mental procedure of checking an email for suspicious features; Rasmussen, 1986).

Optimal, skill-based behaviour involves the automated performance of the task using cue associations (Kahneman & Klein, 2009; Rasmussen, 1986). These associations are formed through repeated exposure to pairings of a situation-specific feature (e.g., a suspicious URL link; Parsons et al., 2016) and an object or event (e.g., phishing email; Klein et al., 2010). Once developed and strengthened, cue associations are stored in long term memory and can then be activated without effort to guide judgements rapidly and unconsciously (Brunswik, 1955; Klein et al., 2010).

At this final stage, an association can be automatically activated (e.g., phishing email) in the presence of an established cue (e.g., a suspicious URL link), and may facilitate the swift identification of a phishing email. In this manner, the process of skill acquisition begins with varied and cumbersome performance, reliant on explicit knowledge and cognitive effort, and progresses to the stage where accurate assessments become unconscious, rapid, and simple (Elliott et al., 2007; Rasmussen, 1986).

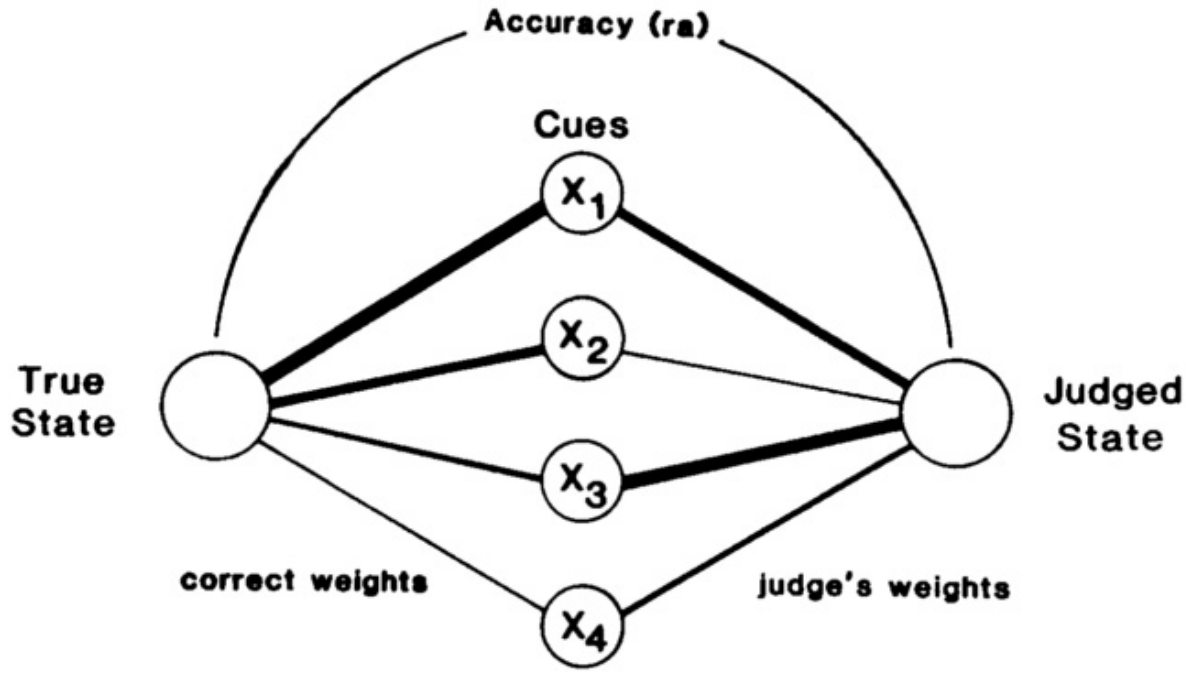
Differences in the capacity to identify and apply the most relevant cues within a given task is referred to as cue utilisation (Loveday et al., 2013; Sturman et al., 2022; Wiggins et al., 2015). Brunswik's (1955) Lens Model provides a framework for how cue utilisation informs judgements about the 'true state' of a situation. It is proposed that the 'true state' (depicted on the left-hand side of Figure 1) is associated with a variety of cues (Brunswik, 1955). In the case of a phishing email (true state), there are a range of environmental cues that are predictive of this true state (e.g., spelling and grammatical errors, suspicious URL link; Parsons et al., 2016).

Depending on the user's level of cue utilisation, the relative importance (weights) of cues used in forming a judgment may be different to those that predict the true state (depicted

on the right-hand side of Figure 1). For example, a user may strongly weight the presence of coloured fonts as evidence that an email is illegitimate, even though this cue may not be predictive of a phishing email. Therefore, a judge may not always utilize the most relevant cues when making an assessment or decision (Brunswik, 1955).

Figure 1

Illustration of Brunswik's Lens Model



Note. Relative weighting of cues indicated by thickness of lines, from Wigton et al. (1986).

There is growing evidence which suggests that cue utilisation may be associated with enhanced phishing email detection (Bayl-Smith et al., 2020; Nasser et al., 2020; Sturman et al., 2022). For instance, Sturman et al. (2022) demonstrated that relatively higher cue utilisation was associated with greater discrimination between genuine and phishing emails. Additionally, there are preliminary indications that cue-based training, which could facilitate cue utilisation, may improve users' capacity for discrimination (Moreno-Fernández et al., 2017; Weaver at al., 2021).

However, the role of cue utilisation in phishing email detection has been examined primarily within samples of undergraduate university students. This population may be more likely to have similar demographic characteristics, cybersecurity related experiences, email habits, and cyber training levels, compared to the broader population (Hanel & Vione, 2016; Sturman et al., 2022). Therefore, the way this specific group interacts with emails may differ from wider society. Thus, it is currently unknown whether greater cue utilisation is also associated with a greater capacity for discrimination within the general population. Research which clarifies the nature of this relationship is critical to inform the development of training programs which are effective for a wide range of users (Sturman et al., 2022).

In addition to knowledge and cue utilisation, another cognitive factor which has been shown to be associated with phishing email detection is the use of different processing modes (Lillie, 2017; Vishwanath, 2015). Dual system theories of processing propose there are two distinct pathways which can be used to make a cognitive judgement (Denes-Raj & Epstein, 1994; Kahneman, 2003). Intuitive processing (System 1; Kahneman, 2003) is automatic, rapid, essentially effortless, and usually the default strategy as it is relatively economical (Denes-Raj & Epstein, 1994). Alternatively, rational processing (System 2; Kahneman, 2003) is characterised by being conscious, analytical, slower, effortful, and intentionally controlled and monitored (Denes-Raj & Epstein, 1994).

Intuitive processing can be limited by its reliance on superficial information (e.g., familiarity with a source; Chen & Chaiken, 1999; Kahneman, 2003) and may hinder effective phishing email detection (Vishwanath, 2015; Vishwanath et al., 2018). Phishing emails aim to elicit intuitive processing to induce a rapid and unconsidered judgement (e.g., to trust the malicious email; Parsons et al., 2019; Wang et al., 2012; Xu & Zhang, 2012). Consequently, individuals who use intuitive processing have been shown to perform worse at detecting phishing emails (Vishwanath, 2015; Vishwanath et al., 2018).

Due to the limitations of intuitive processing, it is often recommended individuals take a rational approach when viewing emails (Khonji et al., 2013; Lillie, 2017). The use of rational processing has been associated with both enhanced phishing email detection and discrimination (Lillie, 2017; Vishwanath, 2015), as well as a range of other factors associated with improved cybersecurity such as reduced trust in emails (Chan-Tin et al., 2021; Gratian et al., 2018; Vishwanath et al., 2018).

Although the literature indicates processing modes are likely to influence phishing email detection, there exist few practical methods of measuring which processing mode is used during a given judgement. However, individuals have differing tendencies to use each mode (Hamilton et al., 2016). Therefore, it may be beneficial to measure users' decision-making styles (Hamilton et al., 2016; Parsons et al., 2019). These are relatively stable, habitual propensities to respond to decision making tasks in a manner which broadly reflects either intuitive (e.g., intuitive style) or rational processing (e.g., rational style; Hamilton et al., 2016).

There is preliminary evidence which suggests decision styles may influence phishing email detection (Chan-Tin et al., 2022; Parsons et al., 2019; Tjostheim & Waterworth, 2020). For instance, Parsons et al. (2019) found an intuitive style predicted worse phishing email detection. However, no research has yet distinguished between detection and discrimination

while examining decision styles. Based on literature which has examined the influence of processing modes (Lillie, 2017; Vishwanath, 2015), it may be logical to suggest a rational style could be associated with both greater detection and discrimination, whereas an intuitive style may be associated with reduced detection and discrimination. However, it remains unclear whether the detection performance outcomes associated with processing are also associated with decision styles.

Additionally, previous studies have used generalised measures of decision styles (Chan-Tin et al., 2022; Tjostheim & Waterworth, 2020). For instance, Parsons et al (2019) utilised the generic Decision Styles Scale (DSS) (e.g., "When making decisions, I rely mainly on my gut feelings"; Hamilton et al. 2016). However, individuals' decision styles differ based on task and environmental factors (Hamilton et al., 2016). Thus, the use of a scale that is contextualised to phishing email detection may increase precision.

The Current Study

Using an age-stratified sample, the present study aims to investigate the roles of phishing email knowledge, cue utilisation, and decision styles (intuitive and rational) in phishing email detection. Using an Email Sorting Task, participants were required to classify a series of 30 emails, half of which were altered to contain features predictive of phishing emails. The Email Sorting Task measured participants' detection performance according to Signal Detection Theory, measuring hits, false alarms, sensitivity (discrimination) and decision criterion (response bias) as dependent variables (Stanislaw & Torodov, 1999). The following hypotheses were presented:

H1:

Participants with greater phishing email knowledge will have more hits and false alarms, no difference in ability to discriminate between genuine and phishing emails, and be

more likely to categorise emails as phishing, compared to their less knowledgeable counterparts.

H2:

Participants with higher cue utilisation will have more hits, less false alarms, a greater ability to discriminate between genuine and phishing emails and be equally as likely to categorise emails as phishing, compared to participants with lower cue utilisation.

H3:

Participants with a greater intuitive decision style will have less hits, more false alarms, a reduced ability to discriminate between genuine and phishing emails and will be equally as likely to categorise emails as phishing, compared to participants with less of an intuitive decision style.

H4:

Participants with a greater rational decision style will have more hits, less false alarms, a greater ability to discriminate between genuine and phishing emails and will be equally as likely to categorise emails as phishing, compared to participants with less of a rational decision style.

Method

Participants

Participants comprised 144 Australian residents (63 males, 81 females) ranging in age from 22 to 86 (M = 52.7, SD = 15.5). Most participants resided in New South Wales (29.8%, n = 43) and Victoria (27.8%, n = 40). The sample was recruited via the Qualtrics Research Panel, an online survey platform, and participants were compensated monetarily for their involvement in the study. Participants reported receiving an average of 18.9 emails per day (SD = 16.6) and spending an average of 2.1 hours per day reading/responding to emails (SD = 2.2). The sample reported spending 6.2 hours per day using a computer (SD = 3.8) and 42.4%

(n = 61) reported being highly confident at using a computer. Most of the sample reported being employed (54.9%, n = 79), primarily in administrative and support services (9%, n = 13) and professional, scientific, and technical services (8%, n = 11).

To participate in the study, participants were required to be over the age of 18, reside in Australia, and able to complete the survey on a laptop or computer. Most participants reported having high levels of motivation (89.6%), engagement (86.9%), effort (88.9%), and attention (77.1%) while completing the Email Sorting Task.

Materials

Participants' ability to detect phishing emails was measured via an Email Sorting Task. Phishing email knowledge was measured using an Objective Phishing Email Knowledge scale that was developed by the researchers of this study. Cue utilisation was measured using the EXPERTise 2.0 (Phishing Edition) battery (Bayl-Smith et al., 2020; Wiggins et al., 2015). Intuitive and rational decision styles were measured using an adapted version of Hamilton et al.'s (2016) Decision Styles Scale. Information security awareness was measured using Parsons et al.'s (2017) HAIS-Q which was included as a covariate within all analyses in this study.

Demographic and Engagement Questions

Participants were asked a range of demographic questions, including age, gender (*Male, Female, Other, Prefer Not to Say*), residential postcode, employment status, and employment industry. Participants were also asked questions relating to email and computer use, including: the number of emails they receive per day (0–100), the number of hours they spend reading/responding to emails (0–24) and using a computer per day (0–24), and their confidence in using a computer (1 = *No Confidence* to 5 = *Very High Confidence*). Following the Email Sorting Task, participants were also asked to rate their agreement with statements

regarding how much effort they put into the task, and how distracted, engaged, and motivated they were during the task (1 = *Strongly Disagree* to 6 = *Strongly Agree*).

Email Sorting Task

Within the Email Sorting Task participants were asked to imagine they were the assistant to Alex Jones, a Professor at The University of Adelaide. In this role, participants had been assigned to sort 30 of Alex Jones' emails into one of 10 pre-determined categories (Urgent, Teaching, Research, Banking, Online Purchases, Social Media Accounts, Official, Spam, Miscellaneous) including Phishing. Of the 30 emails, 15 were genuine and 15 were phishing. Participants were provided with a short description of the criteria for each category (see Appendix A). Phishing was described as "emails which seem fraudulent, fake or otherwise deceptive".

Email Stimuli Development. All the emails used within the Email Sorting Task were genuine emails that had been received by the researchers of this study. Each email contained 100 words or less, and either a URL link in the body of text or an embedded link which could be viewed by hovering the mouse over an image or button (e.g., 'Update Details'; see Figure 3). Fifteen of the 30 emails were systematically manipulated to become phishing emails by including three common phishing email features (see Figure 2): a suspicious URL link (gathered from real phishing emails), an illegitimate sender address (e.g., infoJ44j911F@sculpturalsconces.com), and both a spelling and grammatical error within the first line of included text (e.g., "Thank you for enrolling *too* vote or updating *your,* details."). The remaining 15 emails which had not been altered in any way contained no spelling or grammatical errors and featured both a legitimate URL link and sender's address. These unaltered emails were considered genuine emails (see Figure 3).

Previous literature has identified that phishing emails often utilise a persuasion strategy to encourage a user to click on a malicious URL link (Akbar, 2014; Atkins & Huang,

2013). Some strategies have been found to be more effective than others (Parsons et al., 2019). Consequently, emails either included a common persuasion strategy (authority, scarcity), an uncommon persuasion strategy (social proof, reciprocity), or no persuasion strategy at all (Parsons et al., 2019). Each principle was represented by a total of six emails, where three were genuine and three were phishing emails. The comparison of the effectiveness of these strategies were not under investigation in this thesis but were included to control for the potential influence of these principles.

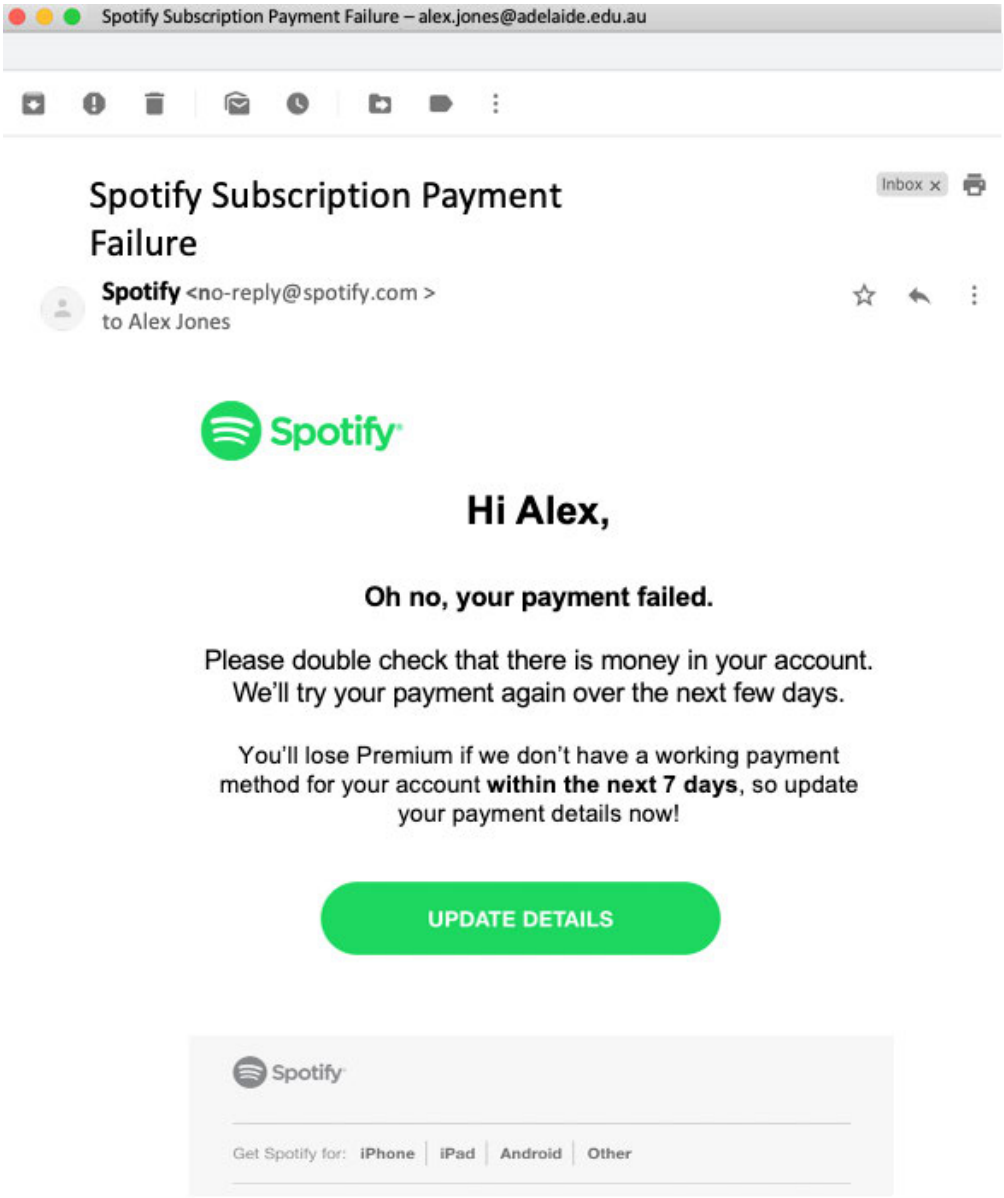
Figure 2

Example of a Phishing Email within the Email Sorting Task

Note. Includes three phishing email features: illegitimate sender address, a spelling and grammatical error within the first line of text, and a suspicious URL link. This email employs the social persuasion principle of authority (Parsons et al., 2019).

Figure 3

Example of a Genuine Email within the Email Sorting Task



Note. The legitimate URL link (https://accounts.spotify.com/en/login) featured in this email is shown when the mouse is hovered over the green "Update Details" button. This email employs the social persuasion principle of scarcity (Parsons et al., 2019).

As part of the task instructions, participants were notified of the ratio of phishing emails present within the task (50%) to control for a shifting response bias. If participants were unaware of the ratio of phishing, some may realise during the task that phishing classifications are relevant to the aims of the study and develop a bias towards classifying emails as phishing (Sturman et al., 2022).

Throughout the task, each email was presented individually, sequentially, and in a randomised order. Each email appeared on screen for 10 seconds (the average duration typically spent viewing an email; Litmus, 2021) before disappearing and prompting the participant to sort the email into a category. This process was repeated until all 30 emails had been viewed and sorted.

EXPERTise 2.0 (Phishing Edition)

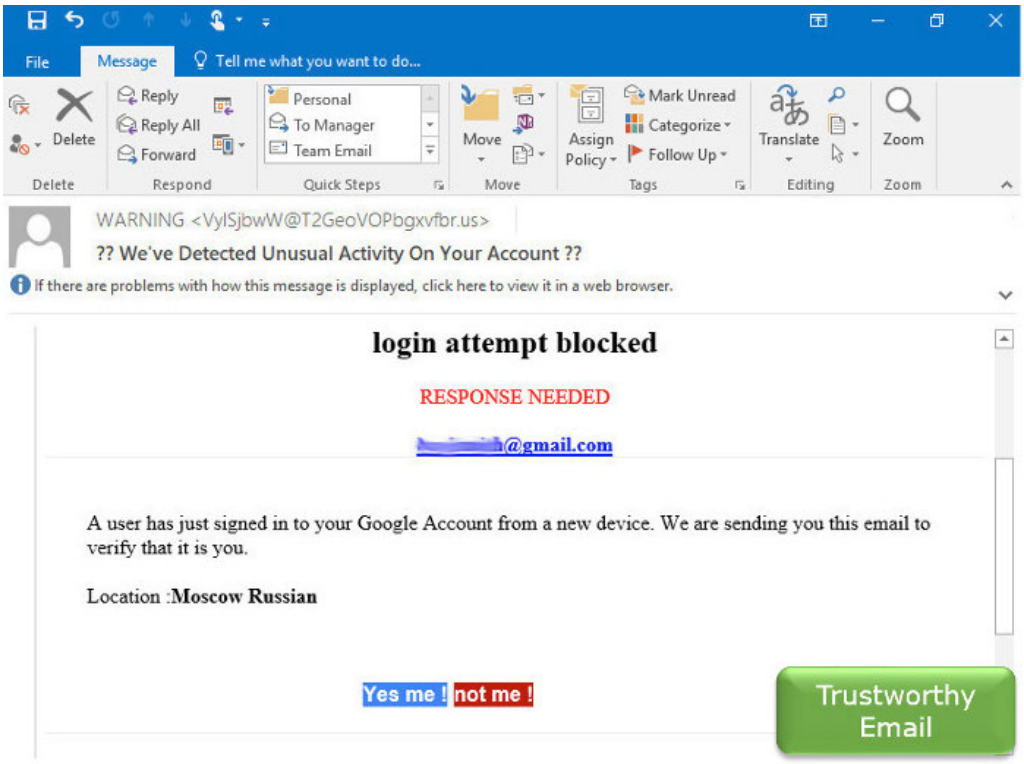
EXPERTise 2.0 in an online shell software tool which measures domain specific cue utilisation (Wiggins et al., 2015). EXPERTise 2.0 has been used to measure cue utilisation in a range of domains, such as phishing email detection (Sturman et al., 2022), medicine (Carrigan et al., 2022), and air traffic control (Falkland & Wiggins, 2019). The tool has shown reasonable predictive validity (Watkinson et al., 2018), construct validity (Wiggins et al., 2014), and test-retest reliability (Loveday et al., 2013). The Phishing Edition of EXPERTise 2.0 was employed in this study and measures the utilisation of cues relevant to the evaluation of the legitimacy of emails (Bayl-Smith et al., 2020). EXPERTise 2.0 comprises four tasks: The Feature Identification Task, the Feature Recognition Task, the Feature Association Task, and the Feature Discrimination Task (Bayl-Smith et al., 2020).

Within the Feature Identification Task, participants were presented with a series of 14 scenarios, each containing either a genuine or phishing email which included features such as a URL link, greeting, sender's address and logo (see Figure 4). For each email, participants were required to click on the area of the email feature which seemed most suspicious, as

quickly as possible. If participants did not believe there were any features of concern within a particular email, they could select an icon titled, “Trustworthy Email” (see Figure 4). The response time between the presentation of the email and the moment the participant selected the area of concern, or “Trustworthy Email”, was measured. Faster response times were associated with higher cue utilisation (Bayl-Smith et al., 2020).

Figure 4

Example of a Phishing Email Presented in Feature Identification Task



During the Feature Recognition Task, participants were shown a series of 20 email stimuli (10 phishing, 10 genuine). Each email was presented for 1000ms, before a new screen prompted participants to classify the email as either “Trustworthy”, “Untrustworthy”, or “Impossible to tell”. The Feature Recognition Task measured the capacity of an individual to rapidly identify the predictive features of phishing emails and utilise them to inform a judgement. Higher accuracy in email classifications was associated with higher cue utilisation (Bayl-Smith et al., 2020; Brouwers et al., 2016).

Within the Feature Association Task, participants were presented with a series of 16 pairs of words associated with phishing (e.g., “Virus”, “Email”), which were presented side by side for 1500ms. After viewing each pair, participants rated their perception of the relatedness of the two words on Likert-scale ranging from 1 (*Extremely unrelated*) to 6 (*Extremely relate*d). The Feature Association Task measured the strength of feature-event relationships in memory which are relevant to phishing emails (Morrison et al., 2013). A greater variance in ratings, in proportion to response time, was associated with higher cue utilisation (Morrison et al., 2013).

During the Feature Discrimination Task, participants were presented with two phishing email scenarios (e.g., an email that claims legal action will be pursued if an unpaid invoice is not settled). Participants were then required to select an appropriate response to each scenario (e.g., “Ignore the email”), before viewing a list of elements contained within the email (e.g., “The sender’s email address”, “Date of email”). Participants then rated the extent to which each feature influenced their initial response to the email on a Likert-scale from 1 (*Not important at all*) to 10 (*Extremely important*). The Feature Discrimination Task measured an individual’s capacity to differentiate between the relative importance of an email’s features when evaluating an email’s legitimacy. Greater variance in ratings of features across both scenarios was associated with higher cue utilisation (Pauley et al., 2009).

Objective Phishing Email Knowledge Scale

Due to the lack of an objective measure of phishing email knowledge available within the literature, a scale was designed for this purpose. The Objective Phishing Email Knowledge scale was created by the researchers to measure participants' knowledge of the predictive features of phishing emails, which can indicate an email is either genuine or phishing. The scale contains 15 statements relating to phishing emails (e.g., "Some phishing emails contain security advice"; see Appendix B; Althobaiti et al., 2021; Lötter & Futcher, 2015; Wang et al., 2012).

The scale has several characteristics which were included to overcome the limitations of previous measures of knowledge. First, it taps knowledge of a variety of relevant predictive features. Second, statements are phrased in a manner which act as an objective test for knowledge. Third, as predictive features can be present in both genuine and phishing emails (Lötter & Futcher, 2015), statements are framed in a hedging manner (e.g., "*Some* phishing emails contain security advice") to ensure consistent interpretation. Fourth, the scale contains six reverse-scored distractor items (see Appendix B). Participants rated the extent to which they agreed with each statement on a 5-point Likert-scale from 1 (*Strongly Disagree*) to 5 (*Strongly Agree*). The responses of the six distractor items were reversed and all responses were summed to create a total score. Higher total scores reflect greater objective phishing email knowledge.

Decision Styles Scale (DSS)

Hamilton et al.'s (2016) Decision Styles Scale (DSS) measured participants' tendency to engage in behaviour which reflects the use of intuitive or rational processing, characterised as intuitive and rational styles. The self-report scale contains 10 statements about how an individual typically makes decisions. The DSS features the two dimensions: intuition (e.g., "When making decisions, I rely mainly on my gut feelings") and rationality (e.g., "I prefer to

gather all the necessary information before committing to a decision") which are each measured by 5-items. Participants rated their agreement with each statement on a 5-point Likert scale ranging from 1 (*Strongly disagree*) to 5 (*Strongly agree*). The responses to items within each domain were summed, with higher scores indicating a greater tendency to use the corresponding decision style. The final scoring outcome results in two independent and separate total scores for intuition and rationality. The DSS has demonstrated high test-retest reliability, and good convergent and divergent validity (Hamilton et al., 2016).

To account for task-dependent nature of decision making (Blais & Weber, 2001; Hamilton et al., 2016), the 10 statements contained in the scale were adapted to fit the context of phishing email detection. The adjusted statements focus on the decision of whether to click on a link within an email. For example, the original statement, "I prefer to gather all the necessary information before committing to a decision" was altered to, "I prefer to gather all the necessary information before *deciding whether to click on a link in an email*" (see Appendix C for full list).

Human Aspects of Information Security Questionnaire (HAIS-Q)

The email-use subscale of the HAIS-Q (Parsons et al., 2014) was used to measure information security awareness as a covariate. The scale features three sub-areas of focus: opening attachments in emails from unknown senders, clicking on links in emails from unknown senders, and clicking on links in emails from known senders (Parsons et al., 2017). Each sub-area is measured by three statements which separately measure participants' knowledge, attitude, and behaviour relating to email-specific information security policy (Parsons et al., 2017).

Participants rated each of the nine statements (e.g., "I am not permitted to click on a link in an email from an unknown sender"; see Appendix D for full list) on a 5-point Likert scale from 1 (*Strongly Disagree*) to 5 (*Strongly Agree*). Five negatively phrased statements

were reverse scored, and responses were summed to create a total score, with higher scores reflecting greater information security awareness. Scores on the email-use subscale have been found to correlate positively with phishing email detection (Parsons et al., 2017) and demonstrates good test-retest reliability (McCormac et al., 2017), internal consistency (α = .78; Parsons et al., 2017) and content validity (Calic et al., 2016; Pattison et al., 2017).

Procedure

This research received ethics approval from the University of Adelaide School of Psychology Institutional Review Board (reference number 20/39). The study was conducted online on the Qualtrics and EXPERTise 2.0 platforms. Participants were selected via Qualtrics to represent a stratified sample based on age. Participants were presented with an information sheet and notified the study was interested in user behaviour and the management of emails. Participants then provided informed consent. They then answered demographic questions, which was followed by the Email Sorting Task. Participants then completed engagement questions, the DSS, HAIS-Q, Objective Phishing Email Knowledge scale, and were then directed to a separate platform to complete the EXPERTise 2.0 (Phishing Edition). The study took participants an average of 38 minutes (*SD* = 21.9) to complete.

Results

Overview of Analyses

Data was analysed in two stages using the IBM Statistical Package for Social Sciences (Version 26). Using a cluster analysis, the first stage involved establishing typologies of cue utilisation (higher, lower) based on performance across the EXPERTise 2.0 tasks. The second stage of analysis examined the hypotheses using analyses of covariance (ANCOVA).

Data Reduction

The data from the EXPERTise 2.0 battery and the Email Sorting Task underwent data reduction. The data reduction for the four EXPERTise 2.0 tasks was consistent with the standard approach used to analyse this data (e.g., Brouwers et al., 2016; Loveday et al., 2013; Sturman et al., 2022). For the Feature Identification Task, the mean response time taken to identify the most suspicious email element across the 14 scenarios was calculated. For the Feature Recognition Task, participants' accuracy across the 20 email classifications was summed. For the Feature Association Task, participants' mean variance in relatedness ratings between the 16 word pairs was calculated, as a proportion of response time. For the Feature Discrimination Task, participants' mean variance in ratings of the importance of the 10 email features across the scenarios was calculated.

Using Signal Detection Theory (Stanislaw & Todorov, 1999), participants received four performance measures for the Email Sorting Task: hit rate, false alarm rate, sensitivity, and decision criterion. A participants' hit rate was the total number of phishing emails that were correctly categorised as phishing. An individual's false alarm rate was the total number of genuine emails that were incorrectly categorised as phishing. Separate *Z* scores were then created for both hit rate and false alarm rate.

Decision criterion and sensitivity were then calculated according to Signal Detection Theory (Stanislaw & Todorov, 1999), consistent with previous phishing detection performance research (Canfield et al., 2016; Sarno & Neider, 2021; Sturman et al., 2022). Higher sensitivity indicates greater discrimination between genuine and phishing emails. A negative decision criterion is associated with conservative responding and a bias towards classifying emails as phishing. Alternatively, a positive decision criterion is associated with liberal responding and a bias towards classifying emails as genuine. The magnitude of decision criterion in either direction reflects the strength of the respective response bias.

Data Analysis

Stage 1: Establishing Typologies

A *k*-means cluster analysis was conducted to categorise participants into two typologies by identifying groups with similar patterns of performance across the four EXPERTise 2.0 tasks (Sturman et al., 2022, Wiggins et al., 2014). The four task scores were converted to *Z* scores prior to the cluster analysis. One participant which had a *Z* score of over 6 on the Feature Identification Task was thus removed from the sample as they were considered an extreme outlier. Consistent with previous research (Bayl-Smith et al., 2020; Sturman et al., 2022), the cluster analysis yielded two typologies which broadly reflected participants with relatively higher (n = 74) and lower (n = 70) levels of cue utilisation. Participants within the higher cue utilisation cluster demonstrated a faster response time on the Feature Identification Task, greater accuracy on the Feature Recognition Task, and greater variance of ratings on both the Feature Association Task and Feature Discrimination Task, relative to participants in the lower cue utilisation cluster (see Table 1). Independent samples t tests indicated statistically significant differences between the two typologies on all four tasks (see Table 1).

Table 1

EXPERTise 2.0 Task Z Scores by Cue Utilisation Typology

EXPERTise 2.0 Tasks	Cue Utilisation Typology	
	Higher (n = 74)	Lower (n = 70)
Feature Identification Task	-0.57**	0.57**
Feature Recognition Task	0.69**	-0.72**
Feature Association Task	0.29**	-0.32**
Feature Discrimination Task	0.17*	-0.21*

Note. *p < .05 (two tailed) **p < .01 (two tailed)

Descriptive Statistics

Descriptive statistics and bivariate correlations for the covariate (HAIS-Q), independent, and dependent variables are reported in Table 2 and Table 3, respectively.

Table 2

Descriptive Statistics

Variable	*M*	*SD*	Min.	Max.
HAIS-Q	34.56	5.99	23	45
Knowledge	60.01	8.22	43	75
Intuitive Style	14.87	5.17	5	25
Rational Style	21.69	3.56	6	25
Hit Rate (*Z* score)	-0.37	1.22	-2.33	2.33
False Alarm Rate (*Z* score)	-0.88	0.84	-2.33	1.11
Sensitivity (*Z* score)	0.51	1.30	-2.58	4.65
Decision Criterion (*Z* score)	0.63	0.83	-1.31	2.33

Table 3

Bivariate Correlations

	Variable	1	2	3	4	5	6	7	8	9
1	HAIS-Q	1								
2	Knowledge	.64**	1							
3	Cue Utilisation	-.08	-.03	1						
4	Intuitive Style	-.37**	-.35**	.08	1					
5	Rational Style	.34**	.49**	.11	.07	1				
6	Hit Rate	.32**	.41**	-.26**	-.30**	.22**	1			
7	False Alarm Rate	.36**	.40**	.10	-.17*	.33**	.25**	1		
8	Sensitivity	.07	.12	-.31**	-.18*	-.01	.78**	-.41**	1	
9	Decision Criterion	-.42**	-.50**	.08	.31**	-.33**	-.87**	-.70**	-.37**	1

Note. *$p < .05$ (two tailed) **$p < .01$ (two tailed)

There was a statistically significant correlation between hit rate and the following variables: false alarm rate, sensitivity, knowledge, cue utilisation, a rational style, and information security awareness. There was a statistically significant negative correlation between hit rate and the following variables: decision criterion and an intuitive style. There was a statistically significant positive correlation between false alarm rate and the following variables: knowledge, a rational style, and information security awareness. There was a statistically significant negative correlation between false alarm rate and the following variables: sensitivity, decision criterion, and an intuitive style. There was a statistically significant negative correlation between sensitivity and the following variables: decision criterion, cue utilisation, and an intuitive style. There was a statistically significant positive correlation between decision criterion and an intuitive style. There was a statistically significant negative correlation between decision criterion and the following variables: knowledge, a rational style, and information security awareness.

Objective Phishing Email Knowledge Scale

As this scale was developed by the researchers for use in this study, analyses were conducted to investigate the psychometric properties of the Objective Phishing Email Knowledge scale. Correlational analyses indicated the scale has good convergent validity, as phishing email knowledge was positively associated with the HAIS-Q. Additionally, the internal reliability of the 15-item knowledge scale was good ($\alpha = .83$).

Stage 2: Hypothesis Testing

To assess the hypotheses, four ANCOVAs were conducted. Each ANCOVA included the four independent variables (knowledge, cue utilisation, and intuitive and rational styles), a covariate (HAIS-Q), and one of the four dependent variables in the model (hit rate, false alarm rate, sensitivity, or decision criterion).

H1. Greater phishing email knowledge was statistically significantly associated with a greater hit rate, $F(1,138) = 6.68$, $p = <.05$, $\eta^2 = .046$ and a more negative decision criterion $F(1,138) = 9.05$, $p = <.05$, $\eta^2 = .062$. There was no relationship found between phishing email knowledge and false alarm rate $F(1,138) = 3.46$, $p = .065$, $\eta^2 = .024$, or sensitivity, $F(1,138) = 1.21$, $p = .273$, $\eta^2 = .009$. As such, H1 was partially supported.

H2. Higher cue utilisation was statistically significantly associated with a greater hit rate $F(1,138) = 10.80$, $p = <.05$, $\eta^2 = .073$ and greater sensitivity $F(1,138) = 13.45$, $p = <.001$, $\eta^2 = .089$, relative to participants with lower cue utilisation. There was no relationship found between relatively higher cue utilisation and false alarm rate $F(1,138) = 1.72$, $p = .192$, $\eta^2 = .012$ or decision criterion $F(1,138) = 3.30$, $p = .072$, $\eta^2 = .023$. Therefore, H2 was fully supported.

H3. A greater intuitive decision style was statistically significantly associated with a lower hit rate $F(1,138) = 4.40$, $p = <.05$, $\eta^2 = .031$, and not with decision criterion $F(1,138) = 3.50$, $p = .63$, $\eta^2 = .025$. Contrary to H3, an intuitive decision style was not statistically significantly associated with false alarm rate $F(1,138) = 0.22$, $p = .68$, $\eta^2 = .002$ or sensitivity $F(1,138) = 2.36$, $p = .497$, $\eta^2 = .003$. Consequently, H3 was partially supported.

H4. A greater rational decision style was not significantly associated with hit rate $F(1,138) = 0.99$, $p = .321$, $\eta^2 = .007$, false alarm rate $F(1,138) = 3.38$, $p = .068$, $\eta^2 = .024$, or sensitivity $F(1,138) = 0.07$, $p = .590$, $\eta^2 = .001$. Additionally, there was no relationship found between a rational decision style and decision criterion $F(1,138) = 3.50$, $p = .064$, $\eta^2 = .025$. Thus, H4 was not supported.

Discussion

The purpose of this study was to gain a better understanding of how knowledge, cue utilisation, and decision styles may influence phishing email detection. In support of H1, participants with greater phishing email knowledge had a greater number of correct phishing

classifications. Further, knowledge was associated with a greater overall propensity for phishing email classifications, but not discrimination, indicating that false alarms increased at a similar rate as hits. Although knowledge was not associated with false alarm rate, this pattern of results suggests participants with greater knowledge were better at detecting phishing emails due to a greater response bias towards classifying emails as phishing.

Additionally, as hypothesised in H2, participants with higher cue utilisation correctly classified more phishing emails compared to those with lower cue utilisation. Furthermore, relatively higher cue utilisation was associated with greater discrimination, but not with a response bias. Contrary to H2, relatively higher cue utilisation was not associated with false rate alarm rate. However, overall, these results indicate that higher cue utilisation was associated with improved phishing email detection due to greater to discrimination between genuine and phishing emails.

Supporting H3, participants with a greater intuitive style demonstrated a reduced ability to correctly classify phishing emails. However, contrary to H3, a greater intuitive style was not associated with false alarms, discrimination, or a response bias. Additionally, contrary to H4, a greater rational style was not associated with detection performance; not hits, false alarms, discrimination, or a response bias.

Consistent with previous literature (Alnajim & Munro, 2009; Kumaraguru et al., 2010; Wang et al., 2012; Zhuo et al., 2022), the results of the present study provide further evidence that phishing email knowledge may enhance phishing email detection. However, consistent with the findings of other researchers (Anandpara et al., 2007; Kumaraguru et al., 2010), the pattern of results also indicates knowledge may improve detection by increasing overall caution towards emails which inadvertently increases the incidence of false alarms. Therefore, possessing greater knowledge of the predictive features of phishing emails may

not be sufficient for individuals to be able to effectively discriminate between phishing and genuine emails.

One possible explanation for this effect may be that although users may have phishing email knowledge, they are not able to effectively apply their knowledge in a practical situation. Users with higher knowledge may have not yet developed the cues required to engage in an optimal, skill-based behaviour. Instead, they may employ knowledge or rule-based behaviour (Rasmussen, 1986). The conscious and effortful application of knowledge associated with these stages of behaviour may require greater cognitive resources (Rasmussen, 1986). These increased demands may conflict with a user's primary task (e.g., sorting emails), reducing their capacity to utilise their knowledge to discriminate between genuine and phishing emails. To compensate, users with greater knowledge may become more conscious of the threat of phishing, which may lead to more cautious email use.

Alternatively, engaging in cue-based processing enables individuals to efficiently make an accurate judgement through the recognition and activation of relevant cues (Brunswik, 1955; Klein et al., 2010; Rasmussen, 1986). Consistent with previous research (Bayl-Smith et al., 2020; Sturman et al., 2022) and Rasmussen's (1986) model of skill acquisition, the present study found that higher cue utilisation, which is associated with skill-based behaviour, corresponded with greater phishing email detection due to an enhanced capacity for discrimination. These findings are consistent with Brunswik's (1955) Lens Model, suggesting that higher cue utilisation enables a judge to more accurately weight and activate the cues most predictive of the 'true state' of an email (e.g., whether it is genuine or phishing). Consequently, these results indicate those who have the capacity to effectively utilise relevant cues may be better able to identify phishing emails without becoming more cautious.

In addition to these findings, the present study extends the understanding of the individual effects of cue utilisation and knowledge on phishing email detection. While previous researchers (Sturman et al., 2022) have demonstrated a relationship between cue utilisation and discrimination beyond information security awareness, the present study shows greater cue utilisation may also enhance discrimination beyond objective phishing email knowledge. Furthermore, the present study found knowledge had a small positive effect on correct phishing email classifications, whereas cue utilisation had a medium effect. This result indicates higher cue utilisation may not only enhance discrimination, but also facilitate a greater degree of detection compared to knowledge. Future research could benefit from investigating cue utilisation and knowledge simultaneously to further understand how these factors may influence detection performance.

Taken together, these findings support the proposition that accurate phishing email detection relies on the identification and utilisation of predictive phishing email features (Grazioli, 2004; Sturman et al., 2022; Wash, 2020). Further, the present study demonstrates that it may not be higher knowledge, but cue utilisation which enables users to effectively use these features to inform an assessment of an email's legitimacy. This finding indicates that the timely and accurate assessment of an email's legitimacy may require automatic, cue-based processing which is reliant on previously established associations (Musuva et al., 2019; Rasmussen, 1986; Sturman et al., 2022). Taken together, these outcomes indicate that increasing users' cue utilisation, in addition to knowledge (Jampen et al., 2020; Zhuo et al., 2022), may be beneficial for promoting greater phishing email detection (Bayl-Smith et al., 2020; Sturman et al., 2022).

Cue-based training has been shown to improve performance in the detection of deception (George et al., 2008) and phishing (Lim et al., 2021; Moreno-Fernández et al., 2017; Weaver et al., 2021). Further, cue-based training which repeatedly exposes participants

to relevant feature-event relationships has been demonstrated to enhance discrimination between genuine and phishing emails (Weaver et al., 2021) and websites (Moreno-Fernández et al., 2017).

The results of the present study build on literature which suggests that the method in which training is delivered may significantly influence its outcomes (Al-Daeef et al., 2017; Kumaraguru et al., 2010; Moreno-Fernández et a., 2017). For instance, the present study indicates that providing users with explicit knowledge of the predictive features of phishing emails (e.g., via security notices or warnings) may be unlikely to improve their discriminative ability (Harrison, 2018; Kumaraguru et al., 2010; Lin et al., 2021). However, providing the same information in a manner which facilitates the development of relevant cues may enhance discrimination (Kumaraguru et al., 2010; Moreno-Fernández et al., 2017; Weaver et al., 2021). Consequently, future research exploring the factors which may influence the effectiveness of knowledge and cue-based training programs may further improve phishing email detection outcomes.

In addition to contributing to a greater understanding of the roles of knowledge and cue utilisation, the present study found mixed results regarding the influence of decision styles. Consistent with Parsons et al. (2019), the results of the present study indicate an intuitive style may be associated with a reduced ability to detect phishing emails. However, the magnitude of this effect was small. Further, the results suggest an intuitive style may not influence discriminative ability, a response bias, or false alarm rate. Therefore, overall, these findings suggest an intuitive style may have little influence in phishing email detection.

Additionally, the findings of the present study regarding the role of a rational style are inconsistent with previous literature which examines the influence of processing (Lillie, 2017; Vishwanath, 2015). Whereas previous research (Lillie, 2017) has shown systematic (rational) processing predicts enhanced discrimination between genuine and phishing emails,

the results of the present study demonstrate that a rational style may not influence detection performance. This outcome, together with the findings of the influence of an intuitive style, suggest that overall decision styles may not be predictive of phishing email detection performance. There are two main factors which may explain the disparity between the outcomes of the present study and previous literature.

First, these contradictory findings may be attributed to differences between decision styles, being habitual tendencies (Hamilton et al., 2016), and the processing used in a specific decision. For instance, previous research (Vishwanath, 2015) has measured which processing mode participants used when opening a single phishing email. It is possible that measuring a specific instance of processing is more precise than measuring a broader pattern of behaviour. For example, whilst a participant of the present study may have exhibited a strong preference for a rational style, and therefore a tendency to engage in rational processing, they may have not necessarily used this type of processing to the same extent within the Email Sorting Task.

In addition to these differences, the measurement of decision styles in the present study was subjective. Participants were asked to consider how they typically make decisions, with no explicit time frame for reference. These factors may have reduced the accuracy of their assessment and elicited a social desirability bias (Nederhof, 1985). For instance, participants may have reported they were more analytical and considerate when using emails than they usually are. To account for these effects, future research could operationalise decision styles by prompting participants on which processing mode was used in a specific series of recent decisions.

Second, it is possible the results of the present study indicate that the processing mode used when viewing emails may have less influence in detection performance than previously suggested (Vishwanath, 2015; Vishwanath et al., 2018). This study was one of the first to consider a range of cognitive factors while measuring decision styles. Thus, previous findings

supporting the effects of processing (Lillie, 2017; Vishwanath, 2015) may have been caused by unmeasured differences of other individual differences. However, given the base of evidence supporting the importance of processing modes in phishing email detection (Lillie, 2017; Vishwanath, 2015; Vishwanath et al., 2018), this interpretation seems unlikely. Nonetheless, future research is warranted to further understand the relationships between processing, decision styles, and detection performance.

Strengths

There are several methodological strengths present within this study. First, this study is one of first to investigate the influence of knowledge, cue utilisation, and decision styles together within a national sample stratified by age. Research which examines the relationship between individual differences and phishing email detection in a sample that reflects broader society may better inform the development of training programs which are effective for a greater number of users. Therefore, it is critical that future research in this area uses samples which are representative of a general population.

Second, this study examined phishing email knowledge using an objective measure. The use of such a measure may facilitate a more precise investigation into the role of knowledge and reduce the impact of extraneous variables such as a social desirability bias or individuals' poor metacognition (Canfield et al., 2019). Future research may benefit from the use of a similarly objective measure.

Third, whereas previous literature has used a generic scale to measure decision styles (Parsons et al., 2019; Tjostheim & Waterworth, 2020), the scale used in the present study was adapted to suit a phishing email detection context. As individuals' decision styles can vary based on task and environmental factors (Hamilton et al., 2016), a scale specified to the demands of the task may be a better reflection of participants' true decision-making

behaviour. Consequently, it is important for future research investigating decision making processes or styles to use similar context-specific measures.

Limitations and Future Directions

There are three main methodological limitations of this study which could be addressed in future research. First, a major limitation was that the experiment was conducted online, as opposed to in a lab or face-to-face setting. Importantly, it is not possible to determine whether participants were providing full attention to the Email Sorting Task. To account for this, participants were prompted to report their attention levels after the task. Although most participants reported paying higher levels of attention, these subjective responses may not necessarily reflect participants' true behaviour. Future research should replicate this study in a controlled environment to improve the reliability and internal validity of results.

Second, although the Email Sorting Task was designed to mimic naturalistic email use, the task may have limited face validity. For example, some participants may never categorise their emails, instead responding exclusively from their inbox. Therefore, the way participants were required to interact with emails within the task may differ from how they typically use emails.

Additionally, in a realistic email setting the prevalence of phishing emails is likely to be substantially lower than the 50% ratio used within the Email Sorting Task (Singh et al., 2019). The elevated ratio used in the present study was selected to account for the possible effect of cognitive fatigue and reduce the overall length of the experiment. To improve the generalisability of research, future studies should prompt participants to engage with emails naturalistically and include a more realistic prevalence of phishing emails.

A final limitation of the Email Sorting Task is that participants dealt with a third party's emails (e.g., "Alex Jones") instead of their own. Individuals are likely to have a

deeper understanding of the types of emails expected to find within their own inbox (Greene et al., 2018), which may be impacted when using someone else's emails. For example, participants may have been unfamiliar with the typical duties or contacts of a professor, and consequently more suspicious of emails with university specific content or terms (e.g., an email from the Vice-Chancellor). Therefore, future research which includes personalised email content may more closely reflect users' detection performance (Greene et al., 2018).

In addition to those already mentioned, the outcomes of this study highlight the importance of one further area of future research. If, as the present study indicates, knowledge-based training may lead to increased caution whereas cue-based training may lead to greater discrimination, then there is a need for further research which evaluates the costs and benefits associated with each approach.

Some researchers (Canfield et al., 2018) have suggested training to increase caution may be more useful than improving discrimination. There is a strong case that many could benefit from engaging with emails more cautiously, particularly the most vulnerable and least knowledgeable individuals (Canfield et al., 2018; Zhuo et al., 2022). However, there are three main limitations that may be associated with increasing caution through knowledge-based training which require further investigation.

First, although increased caution leads to improved detection of phishing emails, it also results in increased false alarms (Parsons et al., 2019). Although missing a genuine email may seem less significant than missing a phishing email, it may still cause considerable consequences. For instance, users may delete or ignore emails which contain opportunities, critical correspondence, fines, or other important messages (Parsons et al., 2019; Sturman et al., 2022). Alternatively, improving discrimination between genuine and phishing emails increases detection performance without the costs of false alarms.

Second, increasing caution alone may not be effective and sustainable. Cybersecurity is typically treated as a secondary aim (Al-Daeef et al., 2017; Zhuo et al., 2022). Therefore, prompting users to become more cautious during email use, without the capacity to accurately identify phishing emails, may lead users to multitask. Multitasking has been associated with a reduced capacity for phishing email detection and is unlikely to be a sustainable strategy to improve phishing email detection in the long term (Kang et al., 2021; Zhuo et al., 2022). In comparison, the activation of cues is effortless and only elicited in the presence of an associated feature (Klein et al., 2010; Lansdale et al., 2010). Thus, cue-based training which increases cue utilisation, and therefore discrimination, may be more likely to improve detection without eliciting multitasking behaviour (Sturman et al., 2019).

Third, research on the retention of the effects of knowledge-based training is mixed (Alnajim & Munro, 2009; Harrison, 2018; Kumaraguru et al., 2010; Mayhorn & Nyeste, 2012). One reason some knowledge-based training programs may have a lower rate of retention is that they may increase caution, which could be challenging to consciously maintain in the long term. Alternatively, as cue associations are stored in long term memory (Klein et al., 2010), cue-based training may be retained more effectively. However, there is distinct lack of literature which examines the retention of cue-based training.

Evidently, much work remains to be done before a clear understanding of the respective costs and benefits of training to increase caution and discrimination is established. Future research should further investigate the overall efficacy, sustainability, and retention of knowledge and cue-based training programs to further improve phishing email detection outcomes.

Conclusion

The aim of the present study was to investigate the roles of knowledge, cue utilisation, and decision styles in phishing email detection. The outcomes of this study

indicate that, while higher phishing email knowledge and cue utilisation may both lead to greater detection of phishing emails, cue utilisation may facilitate greater discrimination, whereas knowledge may not. Further, the findings suggest that higher knowledge may lead users to behave more cautiously, increasing phishing email detection at the cost of increasing false alarms. The results of the present study also indicate that decision styles may not influence detection performance, beyond a small reduction in the ability to detect phishing emails associated with an intuitive style. In an applied context, the outcomes of this study suggest that cue-based training which aims to improve discrimination between genuine and phishing emails, in addition to knowledge-based training, may be beneficial for wider society.

References

Akbar, N. (2014) *Analysing persuasion principles in phishing emails* [Master's Thesis, University of Twente]. University of Twente Student Theses. https://essay.utwente.nl/66177/1/Akbar_MA_EEMCS.pdf

Al-Daeef, M. M., Basir, N., & Saudi, M. M. (2017, 5 July–7 July). *Security Awareness Training: A Review* [Paper presentation]. 2017 World Congress on Engineering, Imperial College, London, United Kingdom.

Alnajim, A., & Munro, M. (2009, 3 April–5 April). *An Evaluation of Users' Anti-Phishing Knowledge Retention* [Paper presentation]. 2009 International Conference on Information Management and Engineering, Kuala Lumpur, Malaysia.

Alseadoon, I., Othman, M. F. I., & Chan, T. (2015). What Is the Influence of Users' Characteristics on Their Ability to Detect Phishing Emails?. In Sulaiman, H., Othman, M., Othman, M., Rahim, Y., Pee, N. (Eds.), *Advanced Computer and Communication Engineering Technology* (vol 315., pp. 949–962). Springer.

Althobaiti, K., Meng, N., & Vaniea, K. (2021, 8 May–13 May). *I don't need an expert! making URL phishing features human comprehensible* [Paper presentation]. 2021 CHI Conference on Human Factors in Computing Systems.

Anandpara, V., Dingman, A., Jakobsson, M., Liu, D., & Roinestad, H. (2007, 12 February–15 February). *Phishing IQ Tests Measure Fear, Not Ability* [Paper presentation]. International Conference on Financial Cryptography and Data Security, Tobago.

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

Atkins, B., & Huang, W. (2013). A study of social engineering in online frauds. *Open Journal of Social Sciences*, *1*(3), 23.

Bayl-Smith, P., Sturman, D., & Wiggins, M. (2020). Cue Utilization, Phishing Feature and Phishing Email Detection. In et al., *Financial Cryptography and Data Security* (FC 2020., pp. 56–70). Springer, Cham.

Blais, A.-R., & Weber, E. U. (2001). Domain-specificity and gender differences in decision making. *Risk, Decision and Policy*, *6*(1), 47–69.

Brouwers, S., Wiggins, M. W., Helton, W., O'Hare, D., & Griffin, B. (2016). Cue utilization and cognitive load in novel task performance. *Frontiers in Psychology*, *7*, 435.

Brunswik, E. (1955). Representative design and probabilistic theory in a functional psychology. *Psychological review*, *62*(3), 193.

Butavicius, M., Parsons, K., Pattinson, M., & McCormac, A. (2016). Breaching the human firewall: Social engineering in phishing and spear-phishing emails. *arXiv preprint arXiv:1606.00887*.

Canfield, C. I., & Fischhoff, B. (2018). Setting priorities in behavioral interventions: An application to reducing phishing risk. *Risk Analysis*, *38*(4), 826–838.

Canfield, C. I., Fischhoff, B., & Davis, A. (2016). Quantifying Phishing Susceptibility for Detection and Behavior Decisions. *Human Factors*, *58*(8), 1158–1172.

Canfield, C. I., Fischhoff, B., & Davis, A. (2019). Better beware: comparing metacognition for phishing and legitimate emails. *Metacognition and Learning*, *14*(3), 343–362.

Caputo, D. D., Pfleeger, S. L., Freeman, J. D., & Johnson, M. E. (2013). Going spear phishing: Exploring embedded training and awareness. *IEEE Security & Privacy*, *12*(1), 28–38.

Carrigan, A. J., Charlton, A., Wiggins, M. W., Georgiou, A., Palmeri, T., & Curby, K. M. (2022). Cue utilisation reduces the impact of response bias in histopathology. *Applied Ergonomics*, *98*, 103590.

Chan-Tin, E., Stalans, L., Johnston, S., Reyes, D., & Kennison, S. (2022). *Predicting Phishing Victimization: Roles of Protective and Vulnerable Strategies and Decision-Making Styles*

[Paper presentation]. Fifth International Workshop on Systems and Network Telemetry and Analytics, New York, NY, USA.

Chaudry, J, A., Chaudry, S, A., Rittenhouse, R, G. (2016). Phishing Attacks and Defences. *International Journal of Security and Its Applications*, *10*(1), 247–256.

Chen, S., & Chaiken, S. (1999). The heuristic-systematic model in its broader context. In Chaiken, H. & Trope, Y. (Eds.), *Dual-process theories in social psychology* (pp. 73–96). The Guilford Press.

Cloudian. (2021, July 15). *Cloudian Ransomware Survey Finds 65% of Victims Pentrated by Phishing had Conducted Anti-Phishing Training* [Press release]. https://cloudian.com/press/cloudian-ransomware-survey-finds-65-percent-of-victims-penetrated-by-phishing-had-conducted-anti-phishing-training/

Australian Competition & Consumer Commission. (2022). *Scamwatch: Scam Statistics*. https://www.scamwatch.gov.au/scam-statistics?scamid=all&date=2021

Darwish, A., El Zarka, A., & Aloul, F. (2012, 18 December–20 December). *Towards understanding phishing victims' profile* [Paper presentation]. 2012 International Conference on Computer Systems and Industrial Informatics, Sharjah, United Arab Emirates.

Das, A., Baki, S., Aassal, A. E., Verma, R., & Dunbar, A. (2020). SoK: A Comprehensive Reexamination of Phishing Research From the Security Perspective. *IEEE Communications Surveys & Tutorials*, *22*(1), 671–708.

Denes-Raj, V., & Epstein, S. (1994). Conflict between intuitive and rational processing: when people behave against their better judgment. *Journal of personality and social psychology*, *66*(5), 819.

Diaz, A., Sherman, A. T., & Joshi, A. (2020). Phishing in an academic community: A study of user susceptibility and behavior. *Cryptologia*, *44*(1), 53–67.

Dou, Z., Khalil, I., Khreishah, A., Al-Fuqaha, A., & Guizani, M. (2017). Systematization of knowledge (sok): A systematic review of software-based web phishing detection. *IEEE Communications Surveys & Tutorials*, *19*(4), 2797–2819.

Downs, J. S., Holbrook, M., & Cranor, L. F. (2007, October). *Behavioral response to phishing risk* [Paper presentation]. 2nd Annual eCrime Researchers Summit, Pittsburgh, Pennsylvania, USA.

Elliott, T., & Mills, V. (2007). Investigating naturalistic decision making in a simulated microworld: What questions should we ask? *Behavior Research Methods*, *39*(4), 901–910.

Falkland, E. C., & Wiggins, M. W. (2019). Cross-task cue utilisation and situational awareness in simulated air traffic control. *Applied Ergonomics*, *74*, 24–30.

George, J. F., Biros, D. P., Burgoon, J. K., Nunamaker Jr, J. F., Crews, J. M., Cao, J., ... & Lin, M. (2008). The role of e-training in protecting information assets against deception attacks. *MIS Quarterly Executive*, *7*(2), 85–97.

Gratian, M., Bandi, S., Cukier, M., Dykstra, J., & Ginther, A. (2018). Correlating human traits and cyber security behavior intentions. *Computers & Security*, *73*, 345–358.

Grazioli, S. (2004). Where did they go wrong? An analysis of the failure of knowledgeable internet consumers to detect deception over the internet. *Group Decision and Negotiation*, *13*(2), 149–172.

Greene, K. K., Steves, M., Theofanos, M. F., & Kostick, J. (2018, 18 February–21 February).). *User context: an explanatory variable in phishing susceptibility* [Paper presentation]. Workshop on Usable Security (USEC) at the Network and Distributed Systems Security (NDSS) Symposium 2018, San Diego, California, USA.

Greitzer, F. L., Li, W., Laskey, K. B., Lee, J., & Purl, J. (2021). Experimental Investigation of Technical and Human Factors Related to Phishing Susceptibility. *Trans. Soc. Comput.*, *4*(2), 1–48.

Hamilton, K., Shih, S.-I., & Mohammed, S. (2016). The Development and Validation of the Rational and Intuitive Decision Styles Scale. *Journal of Personality Assessment*, *98*(5), 523–535.

Hanel, P. H., & Vione, K. C. (2016). Do Student Samples Provide an Accurate Estimate of the General Public? *PloS one*, *11*(12), e0168354.

Harrison, B. (2018). *Does Anti-phishing Training Protect Against Organizational Cyber Attacks?: An Empirical Assessment of Training Methods and Employee Readiness* (Publication No. 10844336) [Doctoral dissertation, State University of New York at Buffalo]. ProQuest One Academic. Ann Arbor. https://www.proquest.com/docview/2226166522?pq-origsite=gscholar&fromopenview=true

Harrison, B., Vishwanath, A., & Rao, R. (2016, 5 January–8 January). *A user-centered approach to phishing susceptibility: The role of a suspicious personality in protecting against phishing* [Paper presentation]. 2016 49th Hawaii International Conference on System Sciences (HICSS), Koala, Hawaii, USA.

Jansen, J., & Leukfeldt, R. (2018). Coping with cybercrime victimization: An exploratory study into impact and change. *Journal of Qualitative Criminal Justice and Criminology*, *6*(2), 205–228.

Jampen, D., Gür, G., Sutter, T., & Tellenbach, B. (2020). Don't click: towards an effective anti-phishing training. A comparative literature review. *Human-centric Computing and Information Sciences*, *10*(1), 33.

Kahneman, D. (2003). A perspective on judgment and choice: mapping bounded rationality. *Am Psychol*, *58*(9), 697–720.

Kahneman, D., & Klein, G. (2009). Conditions for intuitive expertise: A failure to disagree. *American Psychologist*, *64*(6), 515–526.

Kang, M., Shonman, M., Subramanya, A., Zhang, H., Li, X., & Dahbura, A. (2021, 5 January–8 January). *Understanding Security Behavior of Real Users: Analysis of a Phishing Study*

[Paper presentation]. 2021 54th Hawaii International Conference on System Sciences (HICSS), Koala, Hawaii, USA.

Khonji, M., Iraqi, Y., & Jones, A. (2013). Phishing Detection: A Literature Survey. *IEEE Communications Surveys & Tutorials*, *15*(4), 2091–2121.

Klein, G., Calderwood, R., & Clinton-Cirocco, A. (2010). Rapid decision making on the fire ground: The original study plus a postscript. *Journal of Cognitive Engineering and Decision Making*, *4*(3), 186–209.

Kumaraguru, P., Rhee, Y., Sheng, S., Hasan, S., Acquisti, A., Cranor, L. F., & Hong, J. (2007, October). *Getting users to pay attention to anti-phishing education: evaluation of retention and transfer* [Paper presentation]. 2nd Annual eCrime Researchers Summit, Pittsburgh, Pennsylvania, USA.

Kumaraguru, P., Sheng, S., Acquisti, A., Cranor, L. F., & Hong, J. (2010). Teaching Johnny not to fall for phish. *ACM Transactions on Internet Technology (TOIT)*, *10*(2), 1–31.

Lansdale, M., Underwood, G., & Davies, C. (2010). Something Overlooked? How experts in change detection use visual saliency. *Applied Cognitive Psychology*, *24*(2), 213–225.

Lastdrager, E. E. H. (2014). Achieving a consensual definition of phishing based on a systematic review of the literature. *Crime Science*, *3*(1), 9.

Lillie, M. E. (2017). *Think before you click: The effects of systematic processing on phishing susceptibility* [Master's Thesis, The University of Adelaide]. School of Psychology Research & Scholarship. https://digital.library.adelaide.edu.au/dspace/bitstream/2440/131800/1/LillieME_2017_MOHF.pdf

Lim, J., Zhou, L., & Zhang, D. (2021, 2 November–3 November). *Verbal Deception Cue Training for the Detection of Phishing Emails* [Paper presentation]. 2021 IEEE International Conference on Intelligence and Security Informatics (ISI), San Antonio, Texas, USA.

Lin, T., Capecci, D. E., Ellis, D. M., Rocha, H. A., Dommaraju, S., Oliveira, D. S., & Ebner, N. C. (2019). Susceptibility to spear-phishing emails: Effects of internet user demographics and email content. *ACM Transactions on Computer-Human Interaction (TOCHI)*, *26*(5), 1–28.

Litmus. (2021). *State of Email Report*. https://www.litmus.com/resources/state-of-email/

Loveday, T., Wiggins, M., Festa, M., Schell, D., & Twigg, D. (2013). Pattern recognition as an indicator of diagnostic expertise. In P. L. Carmona, J. S. Sanchez & A. L. N. Fred (Eds.), *Pattern recognition-Applications and methods* (Vol. 204, pp. 1–11). Springer.

Lötter, A., & Futcher, L. (2015). A framework to assist email users in the identification of phishing attacks. *Information & Computer Security*, *23*(4), 370–381.

Mayhorn, C. B., & Nyeste, P. G. (2012). Training users to counteract phishing. *Work*, *41*, 3549–3552.

Moreno-Fernández, M. M., Blanco, F., Garaizar, P., & Matute, H. (2017). Fishing for phishers. Improving Internet users' sensitivity to visual deception cues to prevent electronic fraud. *Computers in Human Behavior*, *69*, 421–436.

Morrison, B. W., Wiggins, M. W., Bond, N. W., & Tyler, M. D. (2013). Measuring relative cue strength as a means of validating an inventory of expert offender profiling cues. *Journal of Cognitive Engineering and Decision Making*, *7*(2), 211–226.

Musuva, P. M. W., Getao, K. W., & Chepken, C. K. (2019). A new approach to modelling the effects of cognitive processing and threat detection on phishing susceptibility. *Computers in Human Behavior*, *94*, 154–175.

Nasser, G., Morrison, B. W., Bayl-Smith, P., Taib, R., Gayed, M., & Wiggins, M. W. (2020). The Role of Cue Utilization and Cognitive Load in the Recognition of Phishing Emails. *Frontiers in Big Data*, *3*.

Nederhof, A. J. (1985). Methods of coping with social desirability bias: A review. *European Journal of Social Psychology*, *15*(3), 263–280.

Parsons, K., Butavicius, M., Delfabbro, P., & Lillie, M. (2019). Predicting susceptibility to social influence in phishing emails. *International Journal of Human-Computer Studies*, *128*, 17–26.

Parsons, K., Butavicius, M., Pattinson, M., Calic, D., Mccormac, A., & Jerram, C. (2016). Do users focus on the correct cues to differentiate between phishing and genuine emails? *arXiv preprint arXiv:1605.04717*.

Parsons, K., Calic, D., Pattinson, M., Butavicius, M., McCormac, A., & Zwaans, T. (2017). The Human Aspects of Information Security Questionnaire (HAIS-Q): Two further validation studies. *Computers & Security*, *66*, 40–51.

Parsons, K., McCormac, A., Butavicius, M., Pattinson, M., & Jerram, C. (2014). Determining employee awareness using the Human Aspects of Information Security Questionnaire (HAIS-Q). *Computers & Security*, *42*, 165–176.

Pauley, K., O'Hare, D., & Wiggins, M. (2009). Measuring Expertise in Weather-Related Aeronautical Risk Perception: The Validity of the Cochran–Weiss–Shanteau (CWS) Index. *The International Journal of Aviation Psychology*, *19*(3), 201–216.

Ponemon Institute. (2021). *The 2021 Cost of Phishing Study*. Proofpoint. https://www.proofpoint.com/sites/default/files/analyst-reports/pfpt-us-ar-ponemon-2021-cost-of-phishing-study.pdf

Proofpoint. (2022). *2022 State of The Phish*. Proofpoint. https://www.proofpoint.com/sites/default/files/threat-reports/pfpt-au-tr-state-of-the-phish-2022.pdf

Rasmussen, J. (1986). *Information processing and human-machine interaction. An approach to cognitive engineering*. North-Holland, New York.

Salahdine, F., & Kaabouch, N. (2019). Social Engineering Attacks: A Survey. *Future Internet*, *11*(4), 89.

Sarno, D. M., Lewis, J. E., Bohil, C. J., & Neider, M. B. (2019). Which Phish Is on the Hook? Phishing Vulnerability for Older Versus Younger Adults. *Human Factors*, *62*(5), 704–717.

Sarno, D. M., & Neider, M. B. (2021). So Many Phish, So Little Time: Exploring Email Task Factors and Phishing Susceptibility. *Human Factors*, *0*(0).

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* [Paper presentation]. SIGCHI Conference on Human Factors in Computing Systems, Atlanta, Georgia, USA.

Singh, K., Aggarwal, P., Rajivan, P., & Gonzalez, C. (2019). Training to Detect Phishing Emails: Effects of the Frequency of Experienced Phishing Emails. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *63*(1), 453–457.

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior research methods, instruments, & computers*, *31*(1), 137–149.

Sturman, D., Valenzuela, C., Plate, O., Tanvir, T., Auton, J. C., Bayl-Smith, P., & Wiggins, M. W. (2022). The role of cue utilization in the detection of phishing emails. *Appl Ergon*, *106*, 103887.

Sturman, D., Wiggins, M. W., Auton, J. C., & Loft, S. (2019). Cue utilization differentiates resource allocation during sustained attention simulated rail control tasks. *Journal of Experimental Psychology: Applied*, *25*(3), 317.

Sumner, A., & Yuan, X. (2019, April). *Mitigating phishing attacks: an overview* [Paper presentation]. 2019 ACM Southeast Conference.

Tjostheim, I., & Waterworth, J. A. (2020). Predicting Personal Susceptibility to Phishing. In Rocha, Á., Ferrás, C., Montenegro Marin, C., Medina García, V. (Eds.), Information Technology and Systems. ICITS 2020. Advances in Intelligent Systems and Computing, vol 1137. Springer.

Vishwanath, A. (2015). Examining the Distinct Antecedents of E-Mail Habits and its Influence on the Outcomes of a Phishing Attack. *Journal of Computer-Mediated Communication*, *20*(5), 570–584.

Vishwanath, A., Harrison, B., & Ng, Y. J. (2018). Suspicion, Cognition, and Automaticity Model of Phishing Susceptibility. *Communication Research*, *45*(8), 1146–1166.

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., Herath, T., Chen, R., Vishwanath, A., & Rao, H. R. (2012). Research Article Phishing Susceptibility: An Investigation Into the Processing of a Targeted Spear Phishing Email. *IEEE Transactions on Professional Communication*, *55*(4), 345–362.

Wang, J., Li, Y., & Rao, H. R. (2016). Overconfidence in phishing email detection. *Journal of the Association for Information Systems*, *17*(11), 1.

Wash, R. (2020). How Experts Detect Phishing Scam Emails. *Proc. ACM Hum.-Comput. Interact.*, *4*(CSCW2), Article 160.

Watkinson, J., Bristow, G., Auton, J., McMahon, C. M., & Wiggins, M. W. (2018). Postgraduate training in audiology improves clinicians' audiology-related cue utilisation. *International journal of audiology*, *57*(9), 681–687.

Weaver, B. W., Braly, A. M., & Lane, D. M. (2021). Training Users to Identify Phishing Emails. Journal of Educational Computing Research, 59(6), 1169–1183.

Wiggins, M., Loveday, T., & Auton, J. (2015). EXPERT Intensive Skills Evaluation (EXPERTise) Test. *Sydney: Macquarie University*.

Wiggins, M. W., Azar, D., Hawken, J., Loveday, T., & Newman, D. (2014). Cue-utilisation typologies and pilots' pre-flight and in-flight weather decision-making. *Safety Science*, *65*, 118–124.

Wigton, R. S., Hoellerich, V. L., & Patil, K. D. (1986). How Physicians Use Clinical Information in Diagnosing Pulmonary Embolism: An Application of Conjoint Analysis. *Medical Decision Making*, *6*(1), 2–11.

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

Xu, Z., & Zhang, W. (2012). Victimized by phishing: A heuristic-systematic perspective. *The Journal of Internet Banking and Commerce*, *17*(3), 1–16.

Yang, R., Zheng, K., Wu, B., Li, D., Wang, Z., & Wang, X. (2022). Predicting User Susceptibility to Phishing Based on Multidimensional Features. *Comput Intell Neurosci*, *2022*, 7058972.

Zhuo, S., Biddle, R., Koh, Y. S., Lottridge, D., & Russello, G. (2022). SoK: Human-Centered Phishing Susceptibility. *arXiv preprint arXiv:2202.07905*.

Appendix A

Email Sorting Task Categorisation Options with Descriptions

Which category would you sort this email into?

- ◯ **Urgent** (emails, personal or work-related, that Alex needs to respond to within the next 24-48 hours)
- ◯ **Teaching** (emails from colleagues regarding the coordination of university courses)
- ◯ **Research** (emails regarding Alex's research and research opportunities)
- ◯ **Banking** (Alex's personal banking)
- ◯ **Online purchases** (receipts from purchases Alex has made)
- ◯ **Social Media accounts** (notifications from Alex's social media accounts)
- ◯ **Official** (personal emails from official agencies e.g. Medicare, ATO, AFP)
- ◯ **Spam** (advertisement emails of no consequence)
- ◯ **Phishing** (emails that seem fraudulent, fake or otherwise deceptive)
- ◯ **Miscellaneous** (emails that don't fit into any other category)

Appendix B

Objective Phishing Email Knowledge Scale Items

1. Financial institutions do not usually ask people to disclose passwords over email.
2. An email that contains unrealistic promises can be a sign that it is a phishing email.
3. An email that contains spelling and grammar mistakes can be a sign that it is a phishing email.
4. Even emails that appear to be from people I know can still be phishing emails.
5. An email that contains a generic greeting such as “dear user” can be a sign that it is a phishing email.
6. Some phishing emails use professional branding such as logos and banners
7. Some phishing emails contain security advice.
8. If I receive an email from a company I am affiliated with, it could still be a phishing email.
9. An email that threatens negative consequences can be a sign that it is a phishing email.

Reverse Scored

10. Government institutions often ask for passwords over email.
11. If an email is from someone I know personally, then it is safe to click on any link or disclose any personal information.
12. If an email looks professional (e.g., it contains logos, banners, copyright information) then it is genuine.
13. If an email comes from a company I am familiar with, then it is likely to be genuine.
14. If a hyperlink has the company name in the URL then it is safe to click the hyperlink.
15. An email that does not sign off (e.g., by saying “kind regards”, “cheers”, etc.) can be a sign that it is a phishing email.

Appendix C

Adapted Decision Styles Scale Items

Rational Style Items

1. I prefer to gather all the necessary information before deciding whether to click on a link in an email.
2. I thoroughly evaluate an email before deciding to click on a link in the email.
3. When deciding whether to click on a link in an email, I take time to contemplate the pros/cons or risks/benefits.
4. Investigating the facts is important when deciding whether to click on a link in an email.
5. I weigh a number of different factors when deciding whether to click on a link in an email.

Intuitive Style Items

6. When deciding whether to click on a link in an email, I rely mainly on my gut feelings.
7. My initial hunch about deciding whether to click on a link in an email is generally what I follow.
8. I decide whether to click on links in emails based on intuition.
9. I rely on my first impressions when deciding whether to click on a link in an email.
10. I weigh feelings more than analysis when deciding whether to click on a link in an email.

(Hamilton et al., 2016)

Appendix D

HAIS-Q Email-Use Scale Items

1. I am not permitted to click on a link in an email from an unknown sender.
2. It's risky to open an email attachment from an unknown sender.
3. I don't always click on links in emails just because they come from someone I know.
4. I don't open email attachments if the sender is unknown to me.

Reverse Scored

5. I am allowed to click on any links in emails from people I know.
6. I am allowed to open email attachments from unknown senders.
7. It's always safe to click on links in emails from people I know.
8. Nothing bad can happen if I click on a link in an email from an unknown sender.
9. If an email from an unknown sender looks interesting, I click on the link within it.

(Parsons et al., 2017)