Computers & Security 123 (2022) 102937

Contents lists available at ScienceDirect

Computers & Security

journal homepage: www.elsevier.com/locate/cose

Why people keep falling for phishing scams: The effects of time pressure and deception cues on the detection of phishing emails

Marcus Butavicius[a,*], Ronnie Taib[b], Simon J. Han[b,1]

[a] Department of Defence, Defence Science and Technology Group, 203 LABS, DST Group PO Box 1500, Edinburgh, SA 5111, Australia
[b] Data61, CSIRO, 13 Garden Street, Eveleigh, NSW 2015, Australia

ARTICLE INFO

Article history:
Received 14 October 2021
Revised 21 February 2022
Accepted 25 September 2022
Available online 26 September 2022

Keywords:
Cyber security
Cognitive abilities
Deception
Phishing susceptibility
Decision-making
Habit strength
Signal detection theory
Human-computer interaction

ABSTRACT

Lack of personalisation and poor mechanics (e.g., grammar, spelling and punctuation) are commonly cited as cues of deception that people can use to identify phishing emails. However, in an online email classification experiment ($N = 472$), we found no empirical evidence that the presence of these features was associated with better phishing email discrimination. We also manipulated time pressure and found that it significantly reduced detection accuracy. Participants rarely inspected the URLs associated with links in the phishing emails but, when they did, their detection performance improved. Better performance in distinguishing between genuine and phishing emails was linked to lower levels of an intuitive decision-making style and relatively lower education levels amongst a highly educated sample. Older participants and those with greater computer proficiency and stronger email habit showed a slight increase in tendency to judge emails as suspicious. The results are discussed in terms of intervention strategies such as cyber security training to improve resilience to phishing attacks.

1. Introduction

Phishing emails pose a significant threat to the information security of organisations. In 2020, an estimated 75% of organisations experienced broad-based phishing attacks and 66% were subject to targeted phishing campaigns such as spear-phishing (Proofpoint, 2021). Phishing attacks made up 33% of all cyberattacks, second only to scan-and-exploit attacks (IBM X-Force Incident Response and Intelligence Services, 2021). While technical solutions in the form of email filters exist to detect phishing emails, these are not foolproof (Furnell et al., 2019) and there is evidence that malicious actors are investing in efforts to evade the automated detection techniques through the use of "anti-analysis" features (IBM X-Force Incident Response and Intelligence Services, 2021). As a result, an estimated 74% of US organisations fell victim to at least one phishing attack in 2020 (Proofpoint, 2021). Therefore, people represent the last, and perhaps the most important, line of defence when it comes to thwarting phishing email attacks (Furnell et al., 2006; 2019: Vroom and von Solms, 2004).

As users are the last line of defence, teaching them to identify a phishing email is one of the most important ways an organisation can treat the risk presented by such attacks (Crossler et al., 2013). However, despite efforts to educate users, the phishing email threat is increasing (Furnell et al., 2019) and a recent survey across a variety of organisations showed that 66% of employees clicked on links in phishing emails in simulated phishing exercises (Proofpoint, 2021).

This paper investigates whether signs of deception in phishing emails, known as leakage cues, have any influence on people's resistance to such attacks. We conducted a crowdsourced study focusing on lack of personalisation and poor mechanics, optionally compounded with time pressure as the most representative aspects of modern phishing. This paper extends current research with detailed, multi-faceted, analyses of the results, correlating phishing vulnerability with leakage cues, inspection of hyperlinks, demographics and importantly, decision-making styles and education levels of the participants.

2. Literature review

2.1. Leakage cues

A potentially effective way for users to detect fraudulent emails is to look for the presence of deception in what are known as

* Corresponding author.
E-mail address: marcus.butavicius@dst.defence.gov.au (M. Butavicius).
[1] Present address: Now at Canva, 110 Kippax St, Surry Hills NSW 2010.

https://doi.org/10.1016/j.cose.2022.102937
0167-4048/

'leakage cues' that characterise many phishing emails. Leakage cues are features that uniquely characterise deceptive communication despite the attacker's efforts to appear genuine (Ekman and Friesen, 1969).

Many phishing attacks work by tempting users to click on a link in the email. This is done to direct users to a rogue website in order to glean sensitive information, such as logon credentials and bank account numbers, or to install malicious software (a.k.a. malware) on their device. It is much harder to compromise and modify a legitimate website (i.e., by infecting a genuine website with malware) or to modify the Domain Name Server system (i.e., in order to divert the user from a legitimate site to a malicious one) than it is to simply create a new website. Therefore, most phishing attacks tempt users to click on a link that directs the phish to the hacker's own URL. As a result, the URL is often a good leakage cue of a phishing email, even if the user has to hover over the link in the message to check this. However, previous research has demonstrated that the presence of a suspicious URL does not vastly improve phishing email detection even when all other variables of the email, e.g., social engineering strategy, logo use, personalisation and writing style, are kept constant (Butavicius et al., 2016). This previous study did not capture hover actions, hence it is not known whether the results were caused by not hovering over the link or an inability to interpret the link itself. In the current study, we conducted an email classification task with both genuine and phishing emails via a platform that logged hover actions to answer this question.

In the current study, we also tested the empirical impact of two other leakage cues for phishing emails frequently cited in the literature. These were (1) the use of generic greetings that do not include the recipient's name (Canfield et al., 2016; Egelman et al., 2008; Jakobsson et al. 2007; Parsons et al., 2013); and (2) the presence of poor English mechanics, i.e., mistakes in spelling, grammar and punctuation (Downs et al., 2006; Harrison et al., 2016; Wang et al., 2009, 2012; Wu et al., 2006; Zhou et al., 2004).

Such leakage cues can be the result of the attacker's English skills and access to information sources in creating broad-based phishing attacks. Although attribution of email attacks is always difficult, many attacks on targets in English speaking countries originate in countries where English is not a first language. As a result, such emails often feature failures in English mechanics, i.e., mistakes in spelling, grammar and punctuation, that are a product of a lack of English proficiency and are not always remedied by automated tools such as machine translation. These mechanical errors are often present not just in the email lures created by the attackers but also in the phishing websites linked in these emails (Dhamija et al., 2006). Such mechanical errors may also simply be the result of attackers not belonging to professional organisations (Harrison et al., 2016; Zhou et al., 2004). The mechanical errors may be particularly predictive of deception when the threat actors use the effective social engineering tactic of appealing to authority (Cialdini, 2007) because organisations with authority (e.g., police or taxation office) tend to use highly professional, grammatically correct language in correspondence which will contrast with an imposter's writing style.

Lack of personalisation is also a useful leakage cue as bulk phishing campaigns tend to contain generic greetings (e.g., 'Hi') rather than include the recipient's name (e.g., 'Hi John'). Not all email addresses contain the owner's actual first name in the address themselves and the black-market trade in lists of email addresses do not always contain the owner's given name in the metadata. As a result, greetings in phishing emails sent to these addresses are forced to use a generic greeting. In fact, for phishing email campaigns that are sent to thousands of email addresses, it may not be worth the threat actor's efforts to customise the phishing email with a personalised greeting, even when the first name is contained in their database, because even a very low hit rate amongst the thousands of recipients will be worth the campaign's effort. This contrasts with most legitimate marketing whereby the company has access to the real names of clients and will use this whenever possible to personalise email communication to improve impact and client buy-in as well as to demonstrate their legitimacy, i.e., to verify that they are who they say they are and not a threat actor pretending to be them. This is because a client's records invariably contain metadata such as first names alongside contact information such as email addresses.

However, empirical studies have shown mixed results as to whether users actually pay attention to these leakage cues. On the one hand, studies like Jakobsson et al. (2007) have found that users trusted signs of personalisation when assessing the legitimacy of emails and websites in a lab environment. Similarly, Egelman et al. (2008) found that participants were influenced by personalisation in judging the emails in a phishing detection task and Wang et al. (2012) found that users who self-reported attending to poor grammar in reviewing an email were more likely to correctly identify it as phishing. Molinaro and Bolton (2018) conducted a small (N=10) lab-based study using a role-play scenario where they manipulated cues including poor mechanics and generic greetings. Using Judgement Analysis, both factors featured in the final model, although they were the least predictive of detection accuracy.

On the other hand, there are several studies that contradict these findings and show no evidence that users pay attention to mechanics and personalisation when assessing the legitimacy of emails. In Parsons et al.'s (2015) study, five subject matter experts rated 50 emails in terms of the presence or absence of a number of potential phishing indicators identified in the literature. These judgements indicated that mechanics and personalisation were reliable leakage cues in the emails. However, when average users attempted to classify the same emails in a role play scenario, they ignored such cues and focussed on visceral cues such as legal disclaimers, the quality of visual design and the positive consequences on offer in the email which are unreliable indicators of deception. However, the judgments of the experts were applied post-hoc and, as a result, it was not possible to experimentally control the presentation of leakage cues in the experiment thereby limiting the generalisability of these results.

Similarly, Vishwanath et al. (2011) failed to find evidence for the role of attendance to grammar and spelling cues in an empirical study of 161 participants. Harrison et al., (2016)'s experiment found no evidence of attention and elaboration on poor mechanics on phishing email processing. However, this study only used one email (i.e., did not have a baseline condition where no grammar and spelling errors were present) so was not able to empirically determine the effects of poor mechanics on detection. In Bayl-Smith et al.'s (2020) study using a cue utilisation approach, they found no self-reported evidence that participants were attending to mechanics and personalisation in the detection of a phishing email. In order to examine the role of mechanics and personalisation further, we performed a larger study (N=472) than all the previously cited empirical studies to better examine the influence of these factors. In addition, our study was unique in that we used an experimental methodology that allowed (1) testing of the empirical influence of poor mechanics and personalisation against a baseline (i.e., examining detection performance both in the presence, and absence, of poor mechanics and personalisation); and (2) the application of Signal Detection Theory (Swets, 1964) to more thoroughly investigate the underlying cognitive mechanisms behind email processing. This allowed us to investigate both people's ability to discriminate between phishing and non-phishing emails and their bias towards providing a phishing (or non-phishing) judgment.

2.2. Decision-making style

Previous literature has also found evidence for the influence of decision-making style on phishing detection ability. One popular theory of cognition states that we have two distinct systems for making decisions (Evans and Stanovich, 2013; Petty and Cacioppo, 1986). The first, System 1, involves decision-making processes which are automatic and heuristic-based whereas the second, System 2, uses more effortful, slower and analytic processes. Previous literature suggests that successful phishing email detection may be linked to a bias towards System 2 decision making processes rather than System 1 decision making processes. Several studies have shown that increased impulsivity in decision-making, as measured by Frederick's (2005) Cognitive Reflection Test (CRT), is associated with poorer phishing email detection (Butavicius et al., 2016; Butavicius et al., 2020; Parsons et al., 2013) but not with spear phishing detection (Butavicius et al., 2016). Generally speaking, both attentional and motor impulsivity have been linked with risky cyber security behaviours such as clicking on links in emails from an unknown source (Hadlington, 2017). However, as a general measure of decision-making impulsivity the CRT has several drawbacks: it has been used so often that many participants have already been exposed to it (Thompson and Oppenheimer, 2016), it may tap into numerical ability rather than a general tendency towards analytic reasoning (Welsh, Burns and Delfabbro, 2013) and it can only measure the influence of relative bias between System 1 and 2.

There is also evidence that self-report measures of decision-making preference may predict phishing susceptibility. Vishwanath (2015) demonstrated that phishing susceptibility varied with self-report measures of decision-making preference, i.e., phishing victimisation decreased with both higher levels of systematic processing preference and lower levels of heuristic decision-making preference. Vishwanath's (2015) study adapted previous self-report measures for generic decision-making styles in the extant literature (e.g., Griffin et al. 2002) by focussing only on preferences when processing emails. Parsons et al. (2018) also found that the hit-rate for detecting phishing increased with decreasing scores on Hamilton et al.'s (2016) Intuitive Decision-Making Sub-Scale. However, the results in the literature are not clear cut as Abroshan et al. (2021) failed to find any association between phishing susceptibility on a real-world phishing task and self-reported decision-making style preference as measured by the General Decision-Making Style scale (Scott and Bruce, 1995).

In the current study, we explored the role of analytic and heuristic reasoning in more detail by employing both subscales, Intuitive and Rational, of the self-report Decision Styles Scale questionnaire (Hamilton et al., 2016). This will allow us a more nuanced investigation of decision-making style by examining the separate influence of the two independent components (i.e., rational vs. intuitive) rather than just viewing an individual's style as a single point along a continuum. In contrast to Vishwanath's (2015) paper, we focussed on self-report of generic decision-making preference rather than preference for decision-making specific to emails. This allowed us to see whether phishing susceptibility relates to generic decision-making preference between the two systems or whether it is only associated with email processing.

2.3. Email habit strength

Media habits are formed via the repeated use of a platform and where that repeated use is reinforced under consistent conditions (Verplanken, 2006). For example, regularly checking work email might provide a user with a sense of control over their work tasks. There is some empirical evidence that greater susceptibility to scams on a digital platform may be linked to higher levels of habitual use of that platform. For example, Vishwanath (2014) found that habitual Facebook use, as determined by frequency of use, the size of social networks and inability to self-regulate their use of the platform, was predictive of individual victimisation via social media attacks. Similarly, Vishwanath (2015) also found that individuals were more likely to fall victim to a single phishing email if their email habit tendency, as measured by an adaption of Verplanken and Orbell's (2003) self-report habit index, was stronger. The association of phishing email victimisation with email habit strength was also independent of self-reported preference for systematic or heuristic email processing. While Vishwanath (2015) interpreted this as evidence for habitual processing being independent of System 1 and System 2 processing, Stanovich (2011) has presented evidence that System 1 actually consists of a variety of systems which include both habitual and automated processing. Vishwanath et al. (2018) similarly found an association between high levels of email habit strength and low levels of suspicion both for phishing emails that included an attachment and for those with a hyperlink.

For the current study, we also included the self-report email habit index to investigate the contribution that habit strength has on phishing detection. However, unlike Vishwanath's (2015) 'real' phishing test study, we used a different experimental format, i.e., role play which not only allowed testing on more than one phishing email, it tested processing of genuine emails as well. Testing responses on both genuine and deceitful emails allows a more detailed investigation of the decision-making processes behind email processing using Signal Detection Theory (Swets, 1964). This allowed us to examine the influence of habit strength on both people's ability to distinguish between phishing and genuine emails as well as their bias towards judging an email as malicious.

2.4. Demographics and other variables

We also investigated other factors including the demographic variables of age, gender and education level as well as self-reported measures of proficiency in English and computer usage that may be related to phishing susceptibility. Findings in the previous literature regarding the influence of age and gender are inconsistent and determining their unique impact is difficult due to the number of potential confounding variables (e.g., technical expertise) that are not controlled within, or vary between, different studies (Jampen et al., 2020). Parsons et al. (2018) found evidence that older participants were more accurate in classifying emails but failed to find any evidence of gender effects. However other research has found that females and participants aged below 25 years were more vulnerable to phishing (Jagatic et al., 2007; Sheng et al., 2010), that younger adults were more susceptible but that gender played no role (Sarno et al., 2017) or that age was irrelevant but that females were more susceptible (Abroshan et al., 2021). Taib et al. (2019) found that in a large multinational corporate setting, older participants were more vulnerable than others, as a combination of increased click-through rate (in social proof-based attacks in particular) and lower reporting rate. They found no effect of gender on click-through rates, although female participants were less likely to report than male. Butavicius et al. (2016) found no effects of age or gender on email classification when controlling for the influence of the personality attributes, decision-making style, Information Security Awareness and national culture.

With regards to English proficiency, this variable may be related to the ability to detect phishing emails because better English comprehension skills may help the user to detect the use of poor grammar. In addition, computer proficiency and education levels may be linked to people's ability to judge emails although results in the literature are inconclusive. Parsons et al. (2013) found ev-

idence that people who had completed a course in the area of information security of information systems were actually worse at detecting phishing emails. Pattinson et al. (2015) found that level of education was not linked to Information Security Awareness. In addition, there was a negative correlation with information security awareness and self-reported familiarity with computers, i.e., those with greater familiarity with computers actually engaged in more risky cyber security behaviours. While the results of Pattinson et al. (2015) relate to information security awareness (ISA), performance on ISA tests has previously been strongly associated with performance on phishing tests (Parsons et al., 2017; Butavicius et al., 2020). It is important to investigate the role of computer proficiency to more closely examine the association of phishing detection with email habit strength mentioned above. More specifically, email habit strength may be associated with degree of computer proficiency as both increase with greater usage of computers.

2.5. Time pressure

Finally, we also examined the influence of time pressure on judgements of email safety. From a decision-making perspective, time pressure may degrade a participant's ability to detect phishing by either (1) forcing a judgment to be made before the decision-making process is complete; (2) invoking the faster but more error prone heuristic processes of System 1 in order to respond within the restricted time allowed; or (3) both in combination. Time pressure may also influence motivational factors and affect in ways that are detrimental to cyber security compliance (Chowdhury et al., 2019; 2020). Time pressure has been used as an experimental manipulation to improve motivation and to safely induce a risky environment in studies on phishing website identification (Gopavaram et al., 2021; Kelley and Bertenthal, 2016). However, our literature review found only one study that directly looked at the empirical influence of time pressure on phishing email identification. Jones et al. (2019) found that when users were specifically asked to make quicker responses, they made more mistakes in classifying the legitimacy of emails. In this study, participants were given an overall time limit of 5 min for all tasks, whereas the range of completion times in the baseline condition without a time constraint was 10–15 min. In this present study, we attempted to replicate this finding using a different methodology by imposing a time limit for responding to each individual email to increase the urgency of responses.

3. The current study

Given the potentially important role leakage cues could play in helping people detect phishing emails, and the lack of consistent evidence in the literature that people pay attention to these factors, we conducted an online experiment on email classification that manipulated the use of mechanics and personalisation in phishing attacks. We also examined the influence of time pressure on performance and whether users hovered over the links in emails to inspect the leakage cue of malicious URLs in the task. To the best of our knowledge this is the first experiment that has examined the influence of personalisation, mechanics and time pressure on both phishing and genuine emails which enabled us to apply a Signal Detection Theory approach to more deeply examine the cognitive behaviour involved in processing emails. In order, to examine the association between people's ability to detect phishing emails and individual differences, such as decision-making style, gender, age and proficiency in English and computer use, we included a range of demographic questions and administered a range of additional tests at the completion of the email task itself. In summary, the research questions in the project were:

RQ1 - How frequently do users check destination URLs associated with links in emails and is this hovering behaviour associated with better detection of phishing emails?

RQ2 - Does the presence of either personalisation, poor mechanics or both, predict participants' ability to detect phishing emails?

RQ3 - Does variation in general decision-making preference for either intuitive or rational styles, or both, predict ability to detect phishing emails?

RQ4 - Does email habit strength predict either discrimination, bias, or both in detecting phishing emails?

RQ5 - Does time pressure affect phishing email detection?

RQ6 – How do English proficiency, computer proficiency or demographic variables such as age and gender contribute to phishing susceptibility and / or mediate any of the effects in RQ2- RQ5?

4. Methodology

4.1. Procedure and materials

At the start of the experiment, participants were asked demographic questions (e.g., age range, gender and education level) as well as their self-rated proficiency with computers and English. This was followed by the main email task which used a 'role play' methodology (Parsons et al., 2013). Participants were not explicitly informed they were going to view phishing emails but rather that they were performing an email classification task where emails could be low priority, high priority or suspicious, in order to reduce the subject expectancy effect (Anandpara et al., 2007). The study was approved by CSIRO's human ethics research committee.

In the main task, participants were presented with 37 emails that were modified versions of real emails based on organisations and agencies relevant to the US sample. Each email contained a hyperlink to an external web page that consisted of a generic phrase (e.g., 'Click here' or 'Log in here') that did not reveal the associated URL. Participants were explicitly informed in the instructions that they could hover over the link to reveal the associated URLs, i.e., the website the email would take them to if they clicked on the link. One of these emails was an attention check email that was included to identify participants who were not attending to the task. Twenty-four of the emails were classed as 'genuine' and their URL was legitimate (e.g., https://amazon.com/sites/dejfhle&fe) while for the 12 emails classed as 'phishing' the URL domain was clearly fraudulent and unrelated to the content of the email (e.g., http://sakethounghai.com.vn/plugin/index.htm). The order of emails was a unique random sequence for each participant with the proviso that the attention check email was always included in a random position within the second half of the sequence. To control for effects of persuasive tactics, the emails were designed such that Cialdini's (2007) principles of persuasion were represented equally amongst the 36 emails. Given that not all individuals would have been familiar with the purported sender organisation in each email, participants were asked to assume that each organisation (regardless of their familiarity with it) was relevant to them.

Participants were evenly allocated at random to one of eight experimental conditions using a two by four factor between-participant design. The first factor, called 'Time pressure', contained two levels – a 'Baseline' condition with no time limit and a 'Deadline' condition with an 18 s timeout, a slightly shorter duration than the mean completion time during a pilot study. The second factor, called 'Presentation', contained four conditions:

- *Personalisation*: participants were asked to type in their first name before the start of the task. Each of the genuine emails was modified such that its greeting line contained the name of the participant, e.g., "Hello Phil".
- *Mechanics*: each phishing email was modified to contain spelling and grammar errors.
- *Personalisation + Mechanics:* phishing emails contained the modifications as per the *Mechanics* condition and the genuine emails contained the personalisation featured in the *Personalisation* condition
- *Baseline*: emails were not modified, i.e., genuine emails were not personalised and phishing emails did not include poor spelling and grammar.

For participants in the time pressure condition, i.e., 'Deadline', each email was presented for a maximum of 18 s and a counter was displayed at the top of the screen to indicate the time remaining before automatically moving to the next email. For participants in the 'Baseline' condition, there was no counter nor time-limit to respond.

For each email, participants were asked to classify the email as either 'Urgent', 'Low priority' or 'Suspicious'. At the end of the main email task, participants were asked to complete the 12 item Email Habit Index (Vishwanath, 2015) and then the 10 item decision-making styles questionnaire which consists of five items each for the Rational and Intuitive subscales (Hamilton et al., 2016).

4.2. Participants

542 participants from the US were initially recruited via Amazon's Mechanical Turk but this number was reduced to 472 after removal of those who failed the attention test. There were 176 females (37%) and 1 person who identified as neither male nor female. The most common age bracket was 30-39 (44%) and 154 of the participants (33%) were over the age of 40. The majority rated their computer (60%) and English (82%) proficiency as excellent and education levels were generally high with the majority (58%) having a bachelor's degree and 28% having a post graduate degree.

5. Results

5.1. Effects of conditions

To assess detection accuracy in classifying emails, we collapsed 'Urgent' and 'Low priority' classifications into a single category, reducing the classifications categories to 'Genuine' or 'Suspicious'. Overall accuracy on the email classification task for all email types was only 56% (SD = .144). The average hit rate, i.e., the proportion of phishing emails classified as suspicious, across the experiment was only 42% (SD = .285) and the mean false alarm rate, i.e., the proportion of genuine emails incorrectly classified as suspicious, was 31% (SD = 236). 25 of the participants (5.3%) did not correctly identify a single phishing email throughout the experiment and only 11 (2.6%) identified all 12 phishing emails.

To further analyse results, we used a Signal Detection Theory (SDT) framework (Swets, 1964) to characterise an individual's performance on the email classification task (for more detail see Butavicius et al. 2016). We used A' and B'' which are non-parametric measures of discrimination and bias from the SDT framework, respectively (Stanislaw and Todorov, 1999). By breaking down performance into these two measures, SDT provides better insight into the decision-making process behind the phishing detection task than simple response accuracy. Discrimination refers to an individual's ability to distinguish between a phishing and a genuine email and therefore relates to the quality of our decision-making system, e.g., how well we can correctly identify key phishing cues. A' takes values in the range of 0 to 1 where 1 represents perfect discrimination ability in making classifications while 0.5 represents chance performance. Bias refers to the overall tendency to respond with a 'Genuine' or 'Suspicious' judgment, regardless of the legitimacy of the email being viewed. As such, B'' refers to the threshold at which people make a phishing response irrespective of how well their decision-making system can discriminate between phishing and genuine emails. B'' scores range from -1 (everything is classed as suspicious) to 1 (everything is classed as 'genuine') while 0 represents no bias.

Using these SDT measures, the average ability to differentiate between phishing and genuine emails in the sample was only just above chance (.56, SD = .22) and, on the whole, participants were biased towards classifying an email as genuine (.18, SD = .39). We then conducted a two-way between-participant Analysis of Variance (ANOVA) on A' where the first factor, Text manipulation, had four levels (Baseline, Personalisation, Mechanics and Personalisation + Mechanics) and the second factor, Time pressure, had two levels (Baseline, Time Pressure). Results showed a significant effect of Time pressure ($F(1,448) = 14.982$, $p < .0001$, $\eta^2_p = .032$) but no significant effect of Text manipulation ($p = .319$, $\eta^2_p = .008$) or interaction between the two factors ($p = .775$, $\eta^2_p = .002$). The application of a deadline amounted to a reduction in discrimination ability from .597 (SD=.215) to .5164 (SD = .227). This amounted to a 5.1% reduction in people's accuracy in judging the emails when under the time pressure condition. These trends in A' are displayed in Fig. 1. An equivalent two-way ANOVA on B'' did not yield any significant effects for Text manipulation ($p = .199$, $\eta^2_p = .01$), Time pressure ($p = .944$, $\eta^2_p = < .001$) or their interaction ($p = .863$, $\eta^2_p = .002$).

5.2. Hovering over URLs

As mentioned above, each user was shown 37 emails and each email had a unique hyperlink. However, on average, each user hovered over a link less than 5 times (14% of URLs) in the experiment (Mean = 4.68, SD = 9.55) and 161 of the participants (34% of the sample) did not hover over a single URL in the experiment. The number of hovers was positively correlated with A' ($r = .504$, $p < .001$, $N = 471$).

5.3. Phishing discrimination and other individual characteristics

There were several significant pairwise correlations between the two signal detection theory measures (A' and B'') and variables that measure individual characteristics such as age range, gender, education level, English proficiency, computer proficiency, email habit strength and the intuitive and decision-making style preference scales (see Appendix A). Interestingly, there was no significant correlation between the Intuitive and Rational subscales in this study ($r = -.091$, $N = 447$) nor was there a significant correlation between email habit score and computer proficiency ($r = .013$, $N = 447$). In order to examine the unique influence of each variable on phishing detection, we performed two multiple linear regressions to examine how the variables of age range, gender, intuitive and rational decision-making style preferences, education level, English proficiency, computer proficiency and email habit strength predicted (1) the ability to distinguish between phishing and genuine emails (i.e., A'); and (2) the bias towards declaring a phishing email (i.e., B''). For both regression analyses, there was no evidence of collinearity with Variance Inflation Factors (VIF) for all predictor variables less than 2. In addition, there was no evidence of outliers as Cooks distances were below 1 for all independent variables.

The first linear regression model was significant and accounted for 26% of the variance in A' ($R^2_{adj} = .256$, $F(8,438) = 20.211$, p

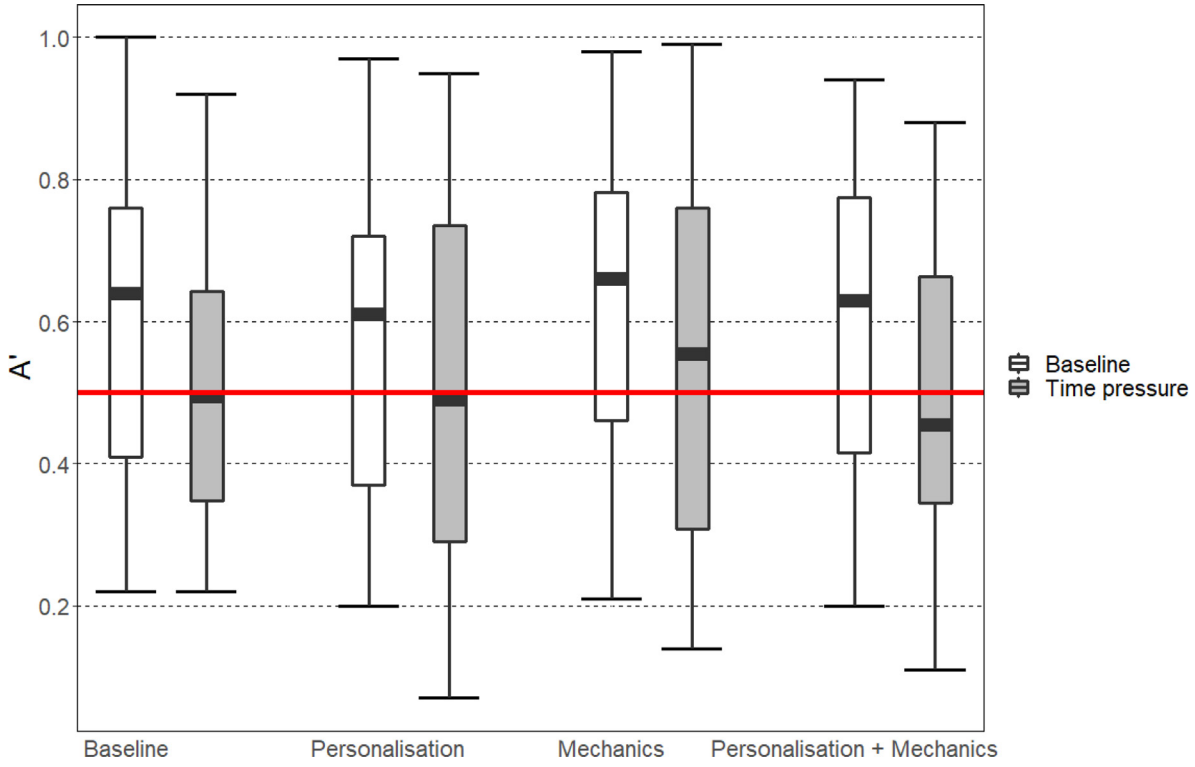

Fig. 1. Boxplots of A' for each treatment condition. The red line indicates chance performance in discrimination, i.e., A' = 0.5.

Table 1
Summary of multiple regression analysis for phishing discrimination (A').

Variable	B	SE B	β Standardised	t	p
Age Range	.01	.08	-.06	5.60	.143
Gender	.01	.02	.02	1.47	.70
Education level	-.03	.01	-.10	.39	.034
English proficiency	0.1	.02	.02	-2.13	.612
Computer proficiency	.01	.02	.04	.51	.406
Email habit score	.02	.02	.06	1.19	.235
Rational subscale	.02	.02	.04	.91	.363
Intuitive subscale	-.11	.01	-.47	-10.50	< .001

Table 2
Summary of multiple regression analysis for phishing discrimination (B").

Variable	B	SE B	β Standardised	t	p
Age Range	-.035	.17	-.10	-2.10	.037
Gender	-.01	.04	-.01	-.26	.797
Education level	-.04	.02	-.08	-1.54	.123
English proficiency	-.04	.04	.06	1.16	.247
Computer proficiency	-.07	.03	-.12	-2.33	.020
Email habit score	-.11	.04	-.15	-2.76	.006
Rational subscale	.01	.04	.02	.32	.746
Intuitive subscale	.02	.02	.04	.87	.385

< .001). The only significant predictors of phishing discrimination ability were the score on the Intuitive Decision-Making subscale (β Standardized = -.467, p < .001) and education level (β Standardized = -.091, p < .001). The second linear regression model was significant but only accounted for 3% of the variance in B" (R^2_{adj} = .03, F(8,438)= 2.72, p = .006). The significant predictors of phishing response bias were the Email Habit score (β Standardized = -.15, p = .006), Computer Proficiency (β Standardized = -.124, p = .02) and Age Range (β Standardized = -.1, p = .037). The results of these two linear regressions are detailed in Tables 1 and 2.

6. Discussion

6.1. Impact of email characteristics on phishing detection

Overall, results showed that people were poor at detecting phishing emails and this was consistent with previous research (Butavicius et al., 2016; Harrison et al., 2016; Parsons et al., 2018; Vishwanath, 2015; Vishwanath et al., 2018). Participants only correctly identified phishing attacks 42% of the time and mistook genuine emails for phishing emails 31% of the time. Our results also showed that individuals did not perform better at detecting phishing emails when either the leakage cues of poor mechanics (i.e., poor grammar, spelling and punctuation), generic greetings, or a combination of the two were present (RQ2). Future research should examine whether this failure was due to inattention to the leakage cues or an inability to interpret them correctly by measuring eye-tracking gaze and / or mouse movement on a similar email processing task (see Yu et al. (2019)).

Very few people regularly checked the URL by hovering over the hyperlink in the email (RQ 1). Before the start of the experiment, participants were informed of this hover functionality but, on average, participants inspected the link only 14% of the time. However, checking the URL was associated with better performance in distinguishing between genuine and fake emails (RQ 1). This suggests that, overall, those who hovered over the link were better able to determine its legitimacy, i.e., they were more likely to pay attention to the URL and correctly analysed it. Further research is needed into how best to help users identify phishing URLs given that even experts can struggle to achieve this without help from third-party tools or services (Althobaiti et al., 2021). Checking the links did not change the bias participants showed in classifying emails, i.e., it did not change people's tendency towards judging emails as legitimate or suspicious overall just because they hovered over the link. Interestingly, this suggests that

hovering over links may be a good behavioural indicator for phishing vigilance that could be monitored via log files in the workplace.

6.2. Impact of personal characteristics on phishing detection

Our results also showed that the only factors that were linked with increased ability to differentiate between phishing and genuine emails were a reduced tendency to prefer an intuitive decision-making style (RQ3) and lower education levels (RQ6). Our findings shed further light on the decision-making processes behind email classification. Previous research has surmised that, because a bias towards rational decision-making style was linked to better phishing detection, that we can improve performance by activating System 2 thinking (Parsons et al., 2017, 2020, 2013; Vishwanath, 2015). The current experiment appears to cast doubt on this conclusion because variation in System 2 thinking did not impact on phishing resilience. Vishwanath (2015) found that better accuracy of detecting phishing emails was linked both to an increase in preference for System 2 and a decrease in preference to System 1 thinking. While we saw a similar trend in the initial pairwise correlations, when we conducted a multifactorial analysis, the effects of a trend in rational thinking appeared to be a second order correlation, i.e., the factor of education explained significantly more unique variance in performance once the covariance with other factors such as rational subscale performance was accounted for. While participants' self-reported tendency towards intuitive decision-making style was linked with poorer accuracy, their preference for a rational mode of thinking did not significantly affect performance. Our results are also consistent with a recent study on phishing susceptibility within social media platforms (Frauenstein and Flowerday, 2020). In this research, susceptibility to fraudulent social media messages was empirically linked to increased preference for heuristic processing but it was independent of variation in preference for systematic processing.

Our findings suggest some independence of the two systems of decision making for processing emails which should be further investigated in studies that can further discriminate different models of engagement of the two systems. For example, rather than just System 1 or System 2 operating exclusively at one time, there are alternative frameworks that propose more complex interactions between the two systems such as parallel (Sloman, 1996) and hybrid (Pennycook et al., 2015) models. Practically speaking, our results would suggest that interventions might be better focussed on strategies to inhibit the use of heuristics rather than trying to activate a more analytic mode of thinking. However, this needs to be investigated empirically because our results showed a trend across participants in a group which may not necessarily translate into viable interventions at the individual level. In general, there is a strong need for such applied research into real-world interventions that could be of practical value in increasing people's resistance to phishing email attacks.

With regards to the influence of education, we interpret our results as being consistent with previous research. Whereas previous research has reported that phishing susceptibility increases with better cyber security expertise, our study found that this susceptibility may be linked to greater levels of general education (and not just in cyber security). Ion et al. (2015) found that cyber security experts (defined as individuals with more than 5 years of experience in cyber security) were more likely to click on links in emails from unknown senders than non-experts. Similarly, whereas Pattinson et al. (2015) and Parsons et al. (2013) showed a link between self-reported risky cyber security behaviours and tertiary education in information systems and information security related subjects, our study showed a link between observed risky cyber security behaviours and higher levels of general education. It should be noted that our sample was, overall, highly educated with over half the participants possessing a bachelor's degree and 28% having a post graduate degree. We speculate that higher levels of education may lead to an overconfidence in the email judgment task. This is consistent with Ion et al. (2015) conclusion that phishing susceptibility is higher amongst cyber security experts because they believed that they "can distinguish between when it's safe and when not to take certain actions" (p.337). Previous research has shown that phishing susceptibility may be linked to overconfidence in external factors such as spam filters (Butavicius et al., 2020), but, to date, the role of overconfidence in innate abilities has not been investigated. Future research should explicitly examine the role of confidence as a variable in predicting phishing email susceptibility.

In our study, the factors of email habit strength, age and computer proficiency were the only variables linked to a bias in judging emails; more specifically, older, more computer proficient participants with a stronger email habit were more likely to judge an email as suspicious (RQ4, RQ6). However, these findings should be interpreted with caution as together these variables only explained a small amount of the variance in the bias of judgments. Our findings do offer a new perspective on how email habit affects phishing email susceptibility first investigated by Vishwanath (2015). By using a SDT framework, we showed that email habit strength was only associated with the decision threshold used by participants to judge an email but not the ability to distinguish between phishing and genuine emails itself. In other words, a stronger email habit was linked to increased bias in classifying emails as suspicious. This contrasts with Vishwanath's (2015) finding that email habit strength was linked to higher rates of falling for phishing emails. Even though Vishwanath (2015) did not use the SDT framework, if email habit strength had been associated with an increase in bias towards classifying emails as suspicious then this would have resulted in higher, not lower, detection rates in their study. The differences in results may be attributable to different experimental methodologies and controlling variables in the two studies. Future research should further examine how email habit strength links to phishing susceptibility.

Our study showed that time pressure significantly increased susceptibility to phishing emails (RQ5). When participants were limited to 18 s to respond to an email, their ability to distinguish between genuine and real emails, A', dropped to nearly chance levels of detection (i.e., 0.52). This amounted in a 5% drop in the accuracy of their email judgments. Our results complement those of Jones et al. (2019) who demonstrated an increase in phishing susceptibility when participants were given an overall time limit to respond to all the emails. This finding demonstrates the increased phishing risk that may be due to real-world factors of deadlines and unreasonable workloads as well as the increased volume of email due to the COVID-19 pandemic (Pathwire, 2021). Future research should look at more realistic experimental manipulations (e.g., time restrictions for processing a group of emails in chronological order and task-based email processing) in order to examine this issue further. In addition, future studies should also examine the effects of varying attention on email processing using dual-task (De Neys, 2006) and cognitive load paradigms (Chen et al., 2016) in order to more closely examine the underlying decision making processes. Such studies should also remain cognisant of alternative theories of decision making to dual systems theories such as those that posit a unified theory of rule-based decision making (Kruglanski and Gigerenzer, 2011).

6.3. Limitations and applications

There are two main limitations of our study. Firstly, our highly educated sample may not be representative of the general public, particularly with regards to the influence of education levels. This

is a problem shared with many other studies on phishing as well as psychology studies in general that have used a sample from a university population (Hanel and Vione, 2016). This should be addressed by more representative samples of the general public in future research. Secondly, we used a role play scenario which lacks the real-world validity of the 'real' phishing emails as users do not actually receive the phishing email in their normal inbox but via an experimental platform.

However, the use of a role-play scenario also provides advantages over the 'real phishing' experiments which may have also explained the different pattern of results we achieved. By including analyses of both phishing and genuine emails, we were able to use a Signal Detection Theory approach which provides a much deeper analysis of the underlying decision-making processes than 'real phishing' methodologies that only collect data on actions performed on the phishing email or studies that only report percentage correct. We also included twelve rather than one phishing email (e.g., Taib 2019) which meant we had a more sensitive measure of phishing susceptibility that also allowed us to control for the effects of different social engineering strategies (Cialdini, 2007). Another reason for the different pattern of results may have been the larger sample size compared to several 'real phishing' studies we have cited, i.e., we had an N of 472 versus 192 (Vishwanath, 2015), 104 and 202 (Vishwanath et al., 2018), and 113 (Harrison et al., 2016). Finally, we included a range of demographic and individual difference variables such as age, gender, education level, computer proficiency which may have controlled for a greater amount of the variance in decision making preference and phishing susceptibility than previous studies. Future work should compare the effectiveness of different phishing email methodologies and examine the use of hybrid approaches that combine processing of different (non-phishing) emails to enable an SDT approach with real-world variables, such as the effects of background information, and other cognitive demands that provide greater face validity with email processing in real-life.

The lack of impact of leakage cues and the general poor performance in classifying emails exhibited in our study paint a depressing picture of people's vulnerability to phishing emails and a failure to educate the general public to mitigate this threat. The fact that users were immune to leakage cues suggests either that a) participants had not received (formal or informal) education about how to detect phishing emails; or b) they had received training but that this had little effect on their ability to detect phishing emails. The quality and content of cyber security training and awareness provided to individuals is crucial given that previous research suggests that many interventions, particularly for phishing, are ineffective. For example, some attempts to educate users, such as through the use of pop-up help files (Jackson et al., 2007) and general phishing training (Anandpara et al., 2007), only increased user's suspicion rather than improving their ability to detect a phishing email. In fact, an empirical study showed that participants were more likely to click on link in a phishing email simply because it contained the word "secure" in the URL (Canova et al., 2014). A literature review on different training techniques for anti-phishing education showed that the most effective training is embedded and personalised (Jampen et al., 2020). Similarly, there is empirical evidence that tailoring cyber security training for individuals, e.g., learning style preference, can significantly improve their level of information security awareness (Pattinson et al., 2019). Our study has highlighted the need for more effective training and awareness campaigns in order to mitigate the continued threat that phishing presents. Such training should be empirically assessed, i.e., such that the effectiveness of different training content and presentation styles should be measured by performance on behavioural tests such as phishing tests and password strength in order to better determine how learning translates into real-world behaviour.

7. Conclusions

As phishing emails remain the attack vector in a third of cybersecurity breaches, and with technical safeguards unable to stop all such attacks, it is vital to understand the human responses to such threats (IBM X-Force Incident Response and Intelligence Services, 2021). Our study showed that signs of deception in phishing emails, known as leakage cues, had no influence on people's resistance to these attacks. More specifically, the presence of leakage cues such as lack of personalisation and poor mechanics was not linked to improved detection of phishing attacks by the general public. In addition, increased time pressure was shown to significantly reduce detection ability in our study. This is particularly concerning given that the volume of emails in the real world continues to increase (Pathwire, 2021). Overall, detection of phishing emails was poor in the study and, while checking URLs in phishing emails was associated with better phishing email detection, this was rarely done. Interestingly, those who had less preference towards making decisions based on intuition and lower levels of education in a highly educated sample were linked to better ability to distinguish between phishing and genuine emails. In addition, those more biased to think that an email was suspicious were older, more proficient in computers and had a stronger habit associated with email use. However, the influence of these variables on bias was small compared with the impact of decision-making style discrimination ability.

The differences with the findings in our study and previous literature were attributed to our use of a role play methodology, control variables, larger sample size and a deeper insight into decision-making provided by our use of Signal Detection Theory. Practically speaking, our results suggest that interventions might better focus on strategies to inhibit the use of heuristics rather than activate a more analytic mode of thinking. Training customised towards personal traits and decision style would set safer email habits and, at the very least, such training should include information on the detection of the leakage cues frequently ignored in our study including URLs, personalisation and mechanics. Future research was proposed into developing better phishing email experimental methodologies and more in-depth studies of the decision-making behind email classification. These will provide insights into not only the theoretical study of the cognitive behaviour behind phishing email processing but also real-world system monitoring and learning interventions.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT authorship contribution statement

Marcus Butavicius: Conceptualization, Methodology, Formal analysis, Writing – original draft, Writing – review & editing, Supervision, Project administration, Funding acquisition. **Ronnie Taib:** Conceptualization, Methodology, Software, Resources, Writing – review & editing, Supervision, Project administration, Funding acquisition. **Simon J. Han:** Methodology, Software, Formal analysis, Investigation, Resources, Data curation, Writing – review & editing.

Acknowledgments

The authors would like to acknowledge the help of Meredith Lillie and Kathryn Parsons in generating the emails for the pilot study cited in this paper and Shlomo Berkovsky, Kun Yu and Fang Chen for early discussions on this topic.

Appendix A. Descriptive statistics and correlation matrix

	1	2	3	4	5	6	7	8	9
1. A'	-								
2. B''	-0.17	-							
3. Age range	0.13*	-0.08	-						
4. Gender	0.10*	-0.01	0.09*	-					
5. Education level	-0.21**	-0.07	-0.10*	-0.08*	-				
6. English proficiency	0.13	-0.01	0.05	0.07	0.06	-			
7. Computer proficiency	0.02	-0.11*	-0.12*	-0.04	0.09*	0.43**	-		
8. Email habit	0.00	-0.15*	-0.01	-0.02	0.05	0.05	0.14*	-	
9. Rational sub-scale	0.13*	-0.06	0.03	0.03	-0.10*	0.14*	0.12*	0.47**	-
10. Intuitive sub-scale	-0.50**	0.00	-0.13*	-0.15*	0.24**	-0.17**	0.04	0.16**	-0.09*
Mean	0.56	0.18	N/A	N/A	4.08	4.77	4.51	3.88	4.13
SD	0.22	0.39	N/A	N/A	0.79	0.54	0.66	0.55	0.55

* $p < .05$, ** $p < .001$.

References

Abroshan, H., Devos, J., Poels, G., Laermans, E., 2021. Phishing happens beyond technology: the effects of human behaviors and demographics on each step of a phishing process. IEEE Access 9, 44928–44949.

Althobaiti, K., Meng, N., Vaniea, K., 2021. I don't need an expert! Making URL phishing features human comprehensible. In: Proceedings of the CHI Conference on Human Factors in Computing Systems, pp. 1–17.

Anandpara, V., Dingman, A., Jakobsson, M., Liu, D., Roinestead, D, 2007. Phishing IQ tests measure fear, not ability. In: Proceedings of the 11th International Conference on Financial Cryptography and 1st International Conference on Usable Security, Heidelberg and Berlin, pp. 362–366.

Bayl-Smith, P., Sturman, D., Wiggins, M., 2020. Cue utilization, phishing feature and phishing email detection. In: Proceedings of the International Conference on Financial Cryptography and Data Security, Cham. Springer, pp. 56–70.

Butavicius, M., Parsons, K., Lillie, M., McCormac, A., Pattinson, M., Calic, D., 2020. When believing in technology leads to poor cyber security: development of a trusty in technical controls scale. Comput. Secur. 98, 102020.

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., Davis, A., 2016. Quantifying phishing susceptibility for detection and behavior decisions. Hum. Factors 58 (8), 1158–1172.

Canova, G., Volkamer, M., Bergmann, C., Borza, R., 2014. NoPhish: an anti-phishing education app. In: Security and Trust Management. Lecture Notes in Computer Science. Springer, Cham, pp. 188–192.

Chen, F., Zhou, J., Wang, Y., Yu, K., Arshad, S.Z., Khawaji, A., Conway, D., 2016. Cognitive load measurement in perspective. In: Robust Multimodal Cognitive Load Measurement. Springer, Cham, pp. 251–254.

Chowdhury, N.H., Adam, M.T., Skinner, G., 2019. The impact of time pressure on cybersecurity behaviour: a systematic literature review. Behav. Inf. Technol. 38 (12), 1290–1308.

Chowdhury, N.H., Adam, M.T., Teubner, T., 2020. Time pressure in human cybersecurity behavior: theoretical framework and countermeasures. Comput. Secur. 97, 101931.

Cialdini, R.B., 2007. Influence: The Psychology of Persuasion. HarperCollins, New York Revised ed.

Crossler, R.E., Johnston, A.C., Lowry, P.B., Hu, Q., 2013. Future directions for behavioural information security research. Comput. Secur. 32 (1), 90–101.

De Neys, W., 2006. Automatic–heuristic and executive–analytic processing during reasoning: chronometric and dual-task considerations. Q. J. Exp. Psychol. 59 (6), 1070–1100.

Dhamija, R., Tygar, J.D., Hearst, M., 2006. Why phishing works. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 581–590.

Downs, J., Holbrook, M.B., Cranor, L.F., 2006. Decision strategies and susceptibility to phishing. In: Proceedings of the 2nd Symposium on Usable Privacy and Security, New York, NY. ACM, pp. 79–90.

Egelman, S., Cranor, L. F., & Hong, J. (2008). You've been warned: an empirical study of the effectiveness of web browser phishing warnings. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (pp. 1065-1074).

Ekman, P., Friesen, W.V., 1969. Nonverbal leakage and cues to deception. Psychiatry 32 (1), 88–105.

Evans, J.B.T., Stanovich, K.E., 2013. Dual-process theories of higher cognition: advancing the debate. Perspect. Psychol. Sci. 8 (3), 223–241.

Frauenstein, E.D., Flowerday, S., 2020. Susceptibility to phishing on social network sites: a personality information processing model. Comput. Secur. 94, 101862.

Frederick, S., 2005. Cognitive reflection and decision making. J. Econ. Perspect. 16 (4), 25–42.

Furnell, S., Jusoh, A., Katsabas, D., 2006. The challenges of understanding and using security: a survey of end-users. Comput. Secur. 25 (1), 27–35.

Furnell, S., Millet, K., Papadaki, M., 2019. Fifteen years of phishing: can technology save us? Comput. Fraud Secur. 2019 (7), 11–16.

Gopavaram, S., Dev, J., Grobler, M., Kim, D., Das, S., Camp, L.J., 2021. Cross-national study on phishing resilience. In: Proceedings of the Workshop on Usable Security and Privacy. USEC.

Griffin, R.J., Neuwirth, K., Giese, J., Dunwoody, S., 2002. Linking the heuristic-systematic model and depth of processing. Commun. Res. 29 (6), 705–732.

Hadlington, L., 2017. Human factors in cybersecurity; examining the link between Internet addiction, impulsivity, attitudes towards cybersecurity, and risky cybersecurity behaviours. Heliyon 3 (7), e00346.

Hanel, P.H., Vione, K.C., 2016. Do student samples provide an accurate estimate of the general public? PloS One 11 (12), e0168354.

Hamilton, K., Shih, S.I., Mohammed, S., 2016. The development and validation of the rational and intuitive decision styles scale. J. Pers. Assess. 98 (5), 523–535.

Harrison, B., Svetieva, E., Vishwanath, A., 2016. Individual processing of phishing emails: how attention and elaboration protect against phishing. Online Inf. Rev. 40 (2), 265–281.

2021 IBM X-Force Incident Response and Intelligence Services (IRIS). 'X-Force Threat Intelligence Index 2021'. IBM Security. Armonk, NY, February 2021. Accessed on 4th August 2021 at https://www.ibm.com/downloads/cas/M1X3B7QG.

Ion, I., Reeder, R., Consolvo, S., 2015. no one can hack my mind": comparing expert and non-expert security practices. In: Proceedings of the 11th Symposium On Usable Privacy and Security. SOUPS 2015, pp. 327–346.

Jackson, C., Simon, D.R., Tan, D.S., Barth, A., Ditrich, S., Dhamija, R.R., 2007. An evaluation of extended validation and picture-in-picture phishing attacks. In: Financial Cryptography and Data Security. Springer Berlin Heidelberg, New York, NY, pp. 281–293.

Jagatic, T.N., Johnson, N.A., Jakobsson, M., Menczer, F., 2007. Social phishing. Commun. ACM 50 (10), 94–100.

Jakobsson, M., Tsow, A., Shah, A., Blevis, E., Lim, Y.K., 2007. What instills trust? A qualitative study of phishing. In: Proceedings of the International Workshop on Usable Security (USEC), Scarborough, Trinidad/Tobago. Springer-Verlag, pp. 356–361.

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

Jones, H.S., Towse, J.N., Race, N., Harrison, T., 2019. Email fraud: the search for psychological predictors of susceptibility. PLoS One 14 (1), e0209684. doi:10.1371/journal.pone.0209684.

Kelley, T., Bertenthal, B.I., 2016. Attention and past behavior, not security knowledge, modulate users' decisions to login to insecure websites. Inf. Comput. Secur. 24 (2), 164–176.

Kruglanski, A.W., Gigerenzer, G., 2011. Intuitive and deliberative judgements are based on common principles. Psychol. Rev. 118, 97–109.

Molinaro, K.A., Bolton, M.L., 2018. Evaluating the applicability of the double system lens model to the analysis of phishing email judgments. Comput. Secur. 77, 128–137.

Parsons, K., Butavicius, M.A., Lillie, M., Calic, D., McCormac, A., Pattinson, M.R., 2018. Which individual, cultural, organisational and interventional factors explain phishing resilience? HAISA 1–11.

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. Comput. Secur. 66, 40–51.

Parsons, K., McCormac, A., Pattinson, M., Butavicius, M., Jerram, C., 2013. Phishing for the truth: a scenario-based experiment of users' behavioural response to emails. In: Security and Privacy Protection in Information Processing Systems -IFIP Advances in Information and Communication Technology. Springer, pp. 366–378.

Pathwire (2021). How the COVID-19 pandemic has changed emailing. Accessed on 16th July 2021 at https://www.mailjet.com/blog/news/covid-19-survey/.

Pattinson, M., Butavicius, M., Lillie, M., Ciccarello, B., Parsons, K., Calic, D., McCormac, A., 2019. Matching training to individual learning styles improves information security awareness. Inf. Comput. Secur. 28 (1), 1–14.

Pattinson, M., Butavicius, M., Parsons, K., McCormac, A., Calic, D., 2015. Factors that influence information security behavior: an Australian web-based study. In: Proceedings of the International Conference on Human Aspects of Information Security, Privacy, and Trust, Cham. Springer, pp. 231–241.

Proofpoint (2021). "2021 State of the Phish: An in-depth look at user awareness, vulnerability and resilience." (Accessed 17 March 2021 2021 at https://www.proofpoint.com/au/resources/threat-reports/state-of-phish).

Parsons, K., Butavicius, M., Pattinson, M., McCormac, A., Calic, D., Jerram, C., 2015. Do users focus on the correct cues to differentiate between phishing and genuine emails? In: Proceedings of the Australian Conference of Information Systems (ACIS). Adelaide December.

Pennycook, G., Fugelsang, J.A., Koehler, D.J., 2015. What makes us think? A three-stage dual-process model of analytic engagement. Cogn. Psychol. 80, 34–72.

Petty, R.E., Cacioppo, J.T., 1986. The elaboration likelihood model of persuasion. In: Communication and Persuasion. Springer, New York, NY, pp. 1–24.

Sarno, D.M., Lewis, J.E., Bohil, C.J., Shoss, M.K., Neider, M.B., 2017. Who are phishers luring?: a demographic analysis of those susceptible to fake emails. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 61, Sage CA: Los Angeles, CA. SAGE Publications, pp. 1735–1739.

Scott, S.G., Bruce, R.A., 1995. Decision-making style: the development and assessment of a new measure. Educ. Psychol. Meas. 55 (5), 818–831.

Sheng, S., Holbrook, M., Kumaraguru, P., Cranor, L., Downs, J., 2010. Who falls for phish? A demo-graphic analysis of phishing susceptibility and effectiveness of

interventions. In: Proceedings of the 28th International Conference on Human Factors in Computing Systems, Atlanta, Georgia, USA, pp. 373–382.
Sloman, S.A., 1996. The empirical case for two systems of reasoning. Psychol. Bull. 119 (1), 3.
Stanislaw, H., Todorov, N., 1999. Calculation of signal detection theory measures. Behav. Res. Methods Instrum. Comput. 31 (1), 137–149.
Stanovich, K.E., 2011. Rationality and The Reflective Mind. Oxford University Press, New York, NY.
Swets, J.A., 1964. Signal Detection and Recognition by Human Observers. Wiley, New York ISBN.
Taib, R., Yu, K., Berkovsky, B., Wiggins, M., Bayl-Smith, P, 2019. Social engineering and organisational dependencies in phishing attacks. In: Lamas, D., Loizides, F., Nacke, L., Petrie, H., Winckler, M., Zaphiris, P. (Eds.), Human-Computer Interaction – INTERACT 2019. Springer International Publishing, pp. 564–584. Lecture Notes in Computer Science.
Verplanken, B., 2006. Beyond frequency: Habit as mental construct. British Journal of Social Psychology 45 (3), 639–656.
Vishwanath, A., 2014. Habitual facebook use and its impact on getting deceived on social media. J. Comput Mediat. Commun. 20, 83–98.
Vishwanath, A., 2015. Examining the distinct antecedents of e-mail habits and its influence on the outcomes of a phishing attack. J. Comput Mediat. Commun. 20 (5), 570–584.
Vishwanath, A., Harrison, B., Ng, Y.J., 2018. Suspicion, cognition, and automaticity model of phishing susceptibility. Commun. Res. 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. Decis. Support Syst. 51, 576–586.
Vroom, C., Von Solms, R, 2004. Towards information security behavioural compliance. Comput. Secur. 23 (3), 191–198.
Wang, J., Chen, R., Herath, T., Rao, H.R, Upadhyaya, S.J., Rao, H.R., 2009. An exploration of the design features of phishing attacks. Annals of Emerging Research in Information Assurance, Security and Privacy Services. Emerald, UK.
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 Trans. Prof. Commun. 55 (4), 345–362.
Wu, M., Miller, R.C., Garfinkel, S., 2006. Do security toolbars actually prevent phishing attacks? In: Proceedings of the Human Factors Computer Systems. Montreal, QC, Canada 2006.
Yu, K., Taib, R., Butavicius, M.A., Parsons, K., Chen, F., 2019. Mouse behavior as an index of phishing awareness. In: Proceedings of the IFIP Conference on Human–Computer Interaction, Cham. Springer, pp. 539–548.
Zhou, L., Burgoon, J.K., Nunamaker, J.F, Twitchell, D., 2004. Automating linguistics-based cues for detecting deception in asynchronous computer-mediated communications. Group Decis. Negot. 13 (1), 81–106.

Marcus Butavicius is a Senior Research Scientist with the Defence Science and Technology Group (DSTG). He joined DSTG in 2001 where he investigated the role of simulation in training, theories of human reasoning and the analysis of biometric technologies. In 2002, he completed a PhD in Psychology at The University of Adelaide on mechanisms of visual object recognition. In 2003 he joined the Intelligence, Surveillance and Reconnaissance Division where his work focused on data visualisation, decision-making, the human aspects of cybersecurity and interface design. He is also a Visiting Research Fellow in the School of Psychology at The University of Adelaide.

Ronnie Taib is a Principal Research Engineer at Data61 – CSIRO in Sydney, Australia. His research revolves around understanding and improving human-computer interaction through behavioural and physiological monitoring. He has focused on cybersecurity for more than six years, exploring the relationship between cognitive load, personality traits, trust and phishing attack vulnerability. These studies range from controlled laboratory conditions to crowdsourcing in a large multinational financial organisation.

Simon J. Han completed his undergraduate studies in psychology and computer science at the University of Melbourne. His research interests lie in cybersecurity and the cognitive sciences. Previously, he was a data analyst at Data61 - CSIRO. He is currently a Machine Learning Engineer at Canva in Sydney, Australia.