

**Behavioural Measures of Phishing Susceptibility:
Examining the Influence of Individual and Situational Factors on Email Management
Decision-Making through Eye-Tracking**

Master Thesis

Jasper Rothert

Psychology of Conflict, Risk & Safety (MSc.)

Faculty of Behavioural, Management and Social Sciences (BMS)

Dr. Steven J. Watson & Dr. Iris van Sintemaartensdijk

University of Twente

June 13th, 2023

Abstract

Phishing attacks continue to pose a significant threat to cybersecurity, with end-users often considered the weakest link. However, it remains unclear why some end-users fall victim to these phishing scams while the large majority does not. This study aims to explore whether cognitive factors and time pressure, which have been linked to increased phishing susceptibility in prior research, affect email management decision-making. Additionally, the study uses eye-tracking technology to investigate the relationship between viewing behaviour to phishing indicators and email judgment performance. By doing so, it seeks to provide a more comprehensive understanding of why certain individuals are more susceptible. The study employs a within-subjects design and recruited 25 participants who completed an email legitimacy task, in which the time to analyze the emails varied halfway through the task. Eye movements were recorded using a Tobii Pro Fusion screen-based eye-tracker. Additionally, participants completed three cognitive tasks assessing working memory capacity, inhibition, and cognitive reflection. The results indicate that time constraints and the cognitive tasks did not impact performance on the email legitimacy task. However, analysis of the eye-tracking data revealed that participants provide the most visual attention (gauged by fixation duration, fixation count, and mean fixation duration) to the sender information of phishing emails, which is positively associated with email judgment performance. Conversely, visual attention to the threat and urgency indicators in phishing emails is negatively associated with judgement performance, and when faced time constraints proportionately more visual attention are given to these indicators compared to not being under time pressure. Visual attention to suspicious URLs in emails shows no clear effect on email judgement performance. This study demonstrates the feasibility of using eye-tracking technology to better understand how individuals visually process phishing emails when making veracity judgments, which could benefit the design of technological and human-centered interventions to mitigate phishing risks.

Behavioural Measures of Phishing Susceptibility: Examining the Influence of Individual and Situational Factors on Email Management Decision-Making through Eye-Tracking.

The rapid growth of the internet has provided numerous benefits for communication, entertainment, and business. However, it has also brought about the serious threat of cybercrime, with approximately 17 percent of the Dutch population, or nearly 2.5 million people, falling victim to cybercrime in 2021 (Centraal Bureau voor de Statistiek, 2022). One specific form of cybercrime is phishing. Phishing is defined as a type of social engineering where attackers, known as phishers, use deceptive tactics to fraudulently obtain sensitive information from unsuspecting users by impersonating trustworthy organizations through electronic communications (Myers, 2007). In 2021, a significant proportion of the Dutch population (68 percent of those aged 15 years and above) reported receiving phishing emails or messages in the past year (Centraal Bureau voor de Statistiek, 2022). Among these individuals, 2 percent admitted to falling for the scams, while 0.8 percent (over 100,000 people) suffered financial losses as a result of phishing. The impact of phishing victimization extends beyond financial consequences, as it can lead to a range of physical, mental, and emotional problems (Button et al., 2014; Coluccia et al., 2020; Ganzini et al., 1990).

Phishing scammers are motivated by various factors, as identified by Weider et al. (2008), including financial gains, identity theft, and even personal fame or notoriety. Although a 2 percent response rate may seem insignificant at first glance, phishing attacks are typically sent in large quantities, reaching thousands of internet users simultaneously (Abu-Nimeh et al., 2007; Garera et al., 2007). Therefore, even with a low response rate, the sheer volume of communications sent makes it economically viable for fraudsters. Furthermore, the widespread sharing of personal information online makes it easier for phishers to gather minimal details about potential victims and craft personalized and believable emails (Vayansky & Kumar, 2018).

Prior research has consistently highlighted human vulnerability as a key factor in phishing susceptibility and overall cybersecurity (Wu et al., 2006; Jones et al., 2015; Jones et al., 2019; Vishwanath et al., 2011; Vishwanath et al., 2018). However, this does not explain why only a small portion of individuals fall victim to phishing schemes while the majority remain unaffected. This raises the question of whether there are systematic differences in susceptibility to phishing victimization at the individual level. Various studies have investigated susceptibility factors, primarily focusing on end-users' decision-making

processes related to email management and considering situational and cognitive influences on these processes (Jones et al., 2015; Jones, 2016).

Often a limitation of phishing susceptibility studies is their reliance on (retrospective) self-report methods to determine the factors influencing individuals' decision to respond or not respond to certain communications. However, these methods are known to be prone to biases (Paulhus & Vazire, 2007; Rosenman et al., 2011). Furthermore, there is currently a lack of research on how people process their decisions regarding email authenticity and what visual information they use during these decision-making processes (McAlanelly & Hills, 2020). Previous studies have used eye-tracking technology to examine how people analyze phishing websites, revealing that the interaction between visual elements of the website and individuals' gaze patterns can predict susceptibility to deception (Miyamoto et al., 2014). Considering the similarities between phishing websites and phishing emails, which both aim to deceive end-users by appearing to be an authentic source, it is plausible that eye-tracking technology could enhance the understanding of email management decision-making as well.

Hence, the primary objective of the present study is to delve deeper into the relationship between viewing behaviour and email management decision-making by employing eye-tracking technology. Additionally, the impact of situational and individual factors, that have been associated with phishing susceptibility in prior research, on email management decision-making will be examined. The situational factor in question involves imposing time constraints during an email legitimacy task, aimed at promoting intuitive decision-making as related to dual process theories. Whilst the individual factors concern cognitive aspects regarding working memory capacity, inhibition, and cognitive reflection.

Characteristics of Fraudulent Emails

Efforts in computer science research aim to effectively detect and prevent phishing emails from reaching end-users' inboxes. However, the effectiveness of these technological countermeasures is often short-lived due to fraudsters adapting their methods to evade new detection techniques (Bergholz et al., 2010; Almomani et al., 2013). Consequently, email service users remain vulnerable to potential attacks, and they bear the responsibility of identifying and handling threats that do manage to bypass the technological defenses (Fette et al., 2007). Detecting phishing communications presents challenges for both humans and technological solutions, primarily because fraudsters can easily create near-identical replicas of trustworthy communications (Fette et al., 2007). Furthermore, fraudsters frequently employ psychological tactics in their communications with the intention of increasing the likelihood

of end-user compliance (Mitnick & Simon, 2002; Atkins & Huang, 2013; Jones et al., 2015; Bullée & Junger, 2020).

Phishing schemes often rely on social engineering, which involves manipulating individuals through psychological tricks to persuade them to assist the attackers (Bullée & Junger, 2020). Phishers attempt to create scenarios that instill enough confidence in the recipients to respond, often using triggers that evoke strong emotions such as dread, excitement, curiosity, or empathy. These triggers ultimately aim to exploit errors in the judgment and decision-making abilities of the recipients (Langenderfer & Shimp, 2001). Theoretical work regarding the psychology of persuasion point to certain techniques that can be used to exploit social norms and could thus increase the probability for people to comply (Cialdini, 2007; Button et al., 2014; Modic & Lea, 2013; Jones et al., 2015). The techniques include for example the use of conformity (mentioning the actions or behaviours of other peers, so that the recipient feels pressured to conform), urgency (the recipient has to respond quickly or something undesirable happens) and threat (the recipient has to respond in order to avoid an undesirable event). Previous research has shown that phishing senders frequently employ these techniques in order to elicit a response (Langenderfer & Shimp, 2001; Atkins & Huang, 2013).

Phishing emails aim to deceive recipients by mimicking communications from trusted sources, making them appear highly authentic at first glance (Fette et al., 2007). Nonetheless, the type of phishing scheme that is used by the fraudster, the quality of the phishing email itself and the inherent nature of how email services function, typically leads to indicators in emails that can be used to help recognizing phishing. The phishing indicators that could be used to help recognizing phishing include for example; the use of social engineering by the sender as discussed above, mismatched or generic email domain names of the sender, and suspicious links embedded in the email. (Langenderfer & Shimp, 2001; Downs et al., 2006; Atkins & Huang, 2013; McAlaney & Hills, 2020). However, as mentioned, the presence of phishing indicators in fraudulent emails is dependent upon multiple factors, meaning that there is no fixed configuration of indicators for all phishing emails.

Situational Susceptibility Factors to Phishing

Previous research on phishing susceptibility has highlighted that the effectiveness of phishing communications goes beyond the quality of the email itself and can be influenced by the situational context in which recipients manage their emails. In real-life situations, end-users of email services are frequently required to manage their emails whilst simultaneously

being pre-occupied with other tasks or working under time constraints (Jones, 2016). These situational factors might affect the ability to make correct judgements and therefore influence the decision-making of end-users (Yan & Gozu, 2012; Jones et al., 2015; Harrison et al., 2016). The Elaboration Likelihood Model of Persuasion (ELM; Petty & Cacioppo, 1986) and the Heuristic-Systematic Model (HSM; Eagly & Chaiken, 1993) explain the different modes of information processing individuals engage in. In the systematic mode of thinking, individuals carefully evaluate available information to form accurate and valid attitudes. However, this mode requires significant cognitive resources, making it impractical to engage in thorough consideration for every decision due to time constraints and limited cognitive capacities. Consequently, individuals are often inclined to form their attitudes in a more simplified manner. In the heuristic mode of thinking, people decide what their attitudes should be using so-called rules of thumb or heuristics, which can introduce biases and lead to errors in judgment (Tversky & Kahneman, 1974).

Applying the ELM and HSM models to the context of phishing suggests that the effectiveness of phishing emails is linked to the thinking mode employed by end-users when processing their emails. Specifically, relying on a heuristic, or intuitive, thinking mode is likely to increase susceptibility to phishing. Previous research provides evidence that when individuals rely on intuitive responses to email stimuli, errors in judging email authenticity are more likely to occur (Yan & Gozu, 2012; Harrison et al., 2016; Jones, 2016). Moreover, the social engineering tactics employed by fraudsters often encourage intuitive thinking, aiming to prompt end-users into making quick and less analytical decisions (Dong et al., 2008; Hadnagy, 2018). An example of this is the use of urgency, in which the fraudster pushes the end-user to provide sensitive information such as their login credentials within 24 hours, or they will lose permanent access to their account. This tactic could trigger a strong emotional response in recipients, such as fear, which promotes less analytical reasoning and prompts them to quickly respond to the perceived threat by complying with the fraudulent request, without thoroughly evaluating the authenticity of the communication. Other psychological models regarding phishing susceptibility have also emphasized the importance of information processing depth and the impact it has on email management decision-making performance (Vishwanath et al., 2011; Vishwanath et al., 2018).

The existing literature highlights that inducing time pressure has a negative impact on email management decision-making performance. However, it remains unclear how time pressure affects the visual processing of electronic communications. Previous studies that have observed a negative effect of time pressure on email management performance

(partially) attributed these results to differences in attitude formation explained by dual process theories such as the ELM and HSM (Yan & Gozu, 2012; Harrison et al., 2016; Jones, 2016). Nonetheless, these findings do not address whether participants in the time pressure condition based their decisions on the same visual information as those without time pressure, or if they processed the visual information differently between the conditions. In other words, it is important to investigate whether time pressure causes individuals to either overlook important indicators of phishing emails or to attend to these indicators of phishing in a different manner compared to when not under time pressure, which ultimately leads to poorer decision-making performance. Miyamoto et al. (2014) suggest that with eye-tracking data, it is possible to predict who would be deceived by a phishing website due to either observing or missing structural elements of the websites that were indicative of phishing. Although this study did not manipulate the time participants had to analyze the stimuli, it did provide a more comprehensive explanation for why people fall for these types of online scams. Therefore, considering that not observing phishing indicators on phishing websites leads to poorer decision-making performance, it is intriguing to investigate whether the same applies to phishing emails. Furthermore, it would be valuable to examine the influence of induced time pressure on individuals' visual processing of emails and its relationship to decision-making performance.

Individual Susceptibility Factors to Phishing

In addition to situational conditions affecting email management decision-making, past research has sought to identify individual factors that influence phishing susceptibility. However, the scientific literature has yielded inconsistent findings regarding variables such as gender, age, and educational background in relation to phishing susceptibility. This inconsistency makes it challenging to pinpoint specific demographic groups that are most vulnerable to victimization (Jagatic et al., 2007; Sheng et al., 2010; Kumaraguru et al., 2010; Darwish et al., 2012; Halevi et al., 2015; Liu et al., 2020; Mohebzada et al., 2012). Another area of investigation has been personality traits and their connection to vulnerability to fraud victimization, using the "Big Five" model of personality. However, the current body of research does not provide reliable predictions of individual susceptibility based on these traits, as the topic lacks an abundance of research and findings have often not been consistently replicated (Modic & Lea, 2012; Workman, 2008; Darwish et al., 2012; Warkentin et al., 2011; Junglas & Spitzmuller, 2006). Therefore, it is necessary to consider other factors that would allow to better differentiate individuals in terms of phishing susceptibility.

Due to the limited research specifically focused on individual differences in phishing susceptibility and online scams, researchers have drawn upon findings from studies related to decision-making, risk behaviour, and consumer behaviour to establish a theoretical foundation in this domain (Jones, 2016). Research regarding phishing susceptibility has attempted to construct a so-called cognitive profile of psychological factors that influence the probabilities of being victimized at the individual level.

Working memory capacity has been linked in prior research with phishing susceptibility (Mayhorn & Nyeste, 2012; Jones et al., 2015). It has been demonstrated that working memory demands vary for a person across different situations and that differences in working memory capacity also exist when comparing individuals (Mayer & Mayer, 2005). Moreover, Cokely and Kelley (2009) found in their study that if participants had an overall higher working memory capacity, they were less likely to engage in risk taking behaviour, compared to individuals with a lower working memory capacity. Applied to the phishing context, this suggests that individuals with lower working memory capacity may be more susceptible to email fraud. Additionally, working memory capacity has also been related to the dual process theories, which suggests that when the situation in which an end-user is managing their emails is sufficiently cognitively demanding, the probability of using an intuitive mode of reasoning to reach a decision increases (Jones et al., 2015). Mayhorn and Nyeste (2012) confirmed in their study that working memory capacity had a significant negative effect on phishing susceptibility. Therefore, the working memory capacity of an individual should be considered as a relevant factor when assessing phishing susceptibility.

Another factor which has been related in research to phishing susceptibility is inhibition. Inhibition can be described as a cognitive process that demands suppressing surrounding information, which then would allow a person to successfully finish a task in question (Conway & Engle, 1994; Redick et al., 2007). When this would be related to email management, it suggests that an increased capacity in inhibition would allow the end-user to suppress their intuitive response and would therefore result to examining all cues available before making an (informed) decision (Jones, 2016). Therefore, inhibition is also theorized to be connected to the dual process theories. Jones (2016) indeed found that inhibitory capacity somewhat predicted the response accuracy in an email legitimacy task by using the Flanker and Stroop test. The same results were observed in the study by Mayhorn and Nyeste (2012), where Stroop scores had a significant negative effect on phishing susceptibility. Thus, inhibition appears to have an influence on phishing susceptibility, and is therefore necessary to include as a predictor.

In previous work, cognitive reflection has been linked to the dual process theories, as well as to phishing susceptibility. People that are high in cognitive reflection are less likely to take a smaller immediate pay out, and instead are more likely to make a calculated risk in order to receive a larger pay out at a later point in time (Frederick, 2005). In context of email management, it would suggest that end-users with a higher cognitive reflection would be more likely to engage in systematic decision-making strategies to weigh the risks and consequences of replying to an email, instead of basing their decisions on intuitive, reward-based responses (Jones, 2016). Jones (2016) indeed found that participants with a higher level of cognitive reflection performed significantly better on a task in which the authenticity of emails had to be determined, as measured by the Cognitive Reflection Task of Frederick (2005). Hence, cognitive reflection is deemed to be a relevant factor when determining phishing susceptibility.

Eye-tracking and Phishing Susceptibility

To fully grasp how an individual engages with email communications it is essential to explore what information and how much of that information is visually processed, which is frequently overlooked in email management decision-making studies (McAlaney & Hills, 2020). One way in which this can be realized is through the use of eye-tracking technology. Simply put, eye-trackers can be used to measure a person's eye movement to determine what they are looking at, which indicates where the point of attention lies (Poole & Ball, 2006; Valtakari et al., 2021). Within the stimulus that the researcher wants the participants to analyze, areas of interest (AOIs) are built in. (Pfeffel et al., 2019). Inside those AOIs, multiple different metrics can be calculated, such as the total duration that an individual fixated their gaze on a specific AOI (fixation duration), the number of times an individual fixated their gaze towards a specific AOI (fixation count) and how long the average fixation lasted in a specific AOI (mean fixation duration). These metrics allow for making inferences regarding the amount of attentional resources an individual is giving to any region of interest of a presented stimulus (Miyamoto et al., 2014; McAlaney & Hills, 2020).

Research specifically focusing on eye-tracking in the context of phishing emails is limited. However, a few studies have been conducted to gain insights into how people analyze electronic communications. Pfeffel et al. (2019) conducted an email legitimacy task with participants who had technical backgrounds in computer sciences. The study observed that participants paid more attention to the top half of the email, where the sender information and the beginning of the body text are located, while allocating less attention to the bottom half of

the email. However, this study used only five broad AOIs within the emails: header, body text, footer, signatures, and attachments. This design limitation makes it difficult to draw inferences about the importance of specific phishing indicators within the emails, as the AOIs for these indicators were not made. For instance, it is unclear whether participants focused on the specific phishing indicator of a wrong domain name in the sender details or other aspects of the sender details, as it was not measured. While Pfeffel et al. (2019) did not draw specific conclusions about phishing indicators and their relevance to phishing susceptibility, it is crucial for the current study to include AOIs for the phishing indicators in order to make meaningful inferences about them.

In a second study examining viewing behaviour and phishing emails, McAlaney and Hills (2020) found no association between the perceived trustworthiness of an email and the total time spent attending to phishing indicators. Participants in this study performed an email legitimacy task where phishing emails contained a single phishing indicator categorized as financial information, urgency, misspellings, or threat. Although participants rated emails that contained phishing indicators as less trustworthy compared to emails without such indicators, the total time spent looking at the phishing indicators did not predict perceived trustworthiness. However, the study did reveal that the phishing indicator AOIs were scanned more intensively and revisited more frequently than would be expected by chance. In other words, participants spent less overall time viewing the phishing indicators than expected by chance, but these indicators required greater attentional resources than expected by chance. Consequently, the precise connection between viewing behaviour, phishing indicators, and email management decision-making remains unclear and might be more complex than expected. Nevertheless, the study provided evidence that eye-tracking can be used to determine whether individuals look at phishing indicators and the order in which they attend to them.

The previous studies mentioned have not examined whether the allocation of attentional resources to phishing indicators is associated with performance on an email management decision-making task. Furthermore, they did not investigate whether the allocation of attentional resources to these indicators varies depending on the presence of time constraints (situational context). Exploring these aspects could provide more insights into why some individuals are deceived by phishing emails while others are not.

The current study

This study aims to provide a deeper understanding of how individuals visually analyze authentic and fraudulent emails in an email legitimacy task and how this affects decision-making. Additionally, the study will manipulate the time available for participants to analyze the emails, with one condition imposing time constraints and the other condition allowing sufficient time. Furthermore, the study will measure three cognitive factors - working memory capacity, inhibition, and cognitive reflection - and examine their relationship with email management decision-making performance. Based on these objectives, the following hypotheses have been formulated:

Hypothesis 1: Participants in the induced time pressure condition will perform with lower accuracy on the email legitimacy task.

Hypothesis 2: Participants with a lower working memory capacity will perform with lower accuracy on the email legitimacy task.

Hypothesis 3: Participants with lower inhibitory capacity will perform with lower accuracy on the email legitimacy task.

Hypothesis 4: Participants with lower cognitive reflection will perform with lower accuracy on the email legitimacy task.

Hypothesis 5: Participants in the induced time pressure condition will provide proportionately less attentional resources to the phishing indicators.

Hypothesis 6: The amount of attentional resources given to the phishing indicators is associated with judgement performance on the email legitimacy task.

Methods

Participants

Psychology and communication science students from the University of Twente were recruited through voluntary response sampling, in which participants enrolled themselves through the University's Sona system to participate in the study in exchange for course credits, as well as through convenience sampling. The inclusion criterium was that participants had normal or corrected-to-normal vision. An exclusion criterium that prohibited participation was if the screen-based eye-tracker was not able to calibrate according to the manufacturer's standardized procedure. This resulted in a sample of 25 participants which comprised 14 females and 11 males. Participants ranged in age from 18 to 26, with a mean age of 21.20 ($SD = 2.10$). The nationality of the participants was German ($N = 11$), Dutch ($N = 10$) and a smaller proportion of other nations which included: Italian, Chinese, Vietnamese and Kazakh ($N = 4$).

Design

A within-subjects design was employed with a manipulation to time pressure (time pressure vs. no time pressure) halfway through an email legitimacy task. In the email legitimacy task, the participants had to assess the authenticity of 32 electronic communications and their eye-movement was tracked using a screen-based eye-tracker. Of these 32 email stimuli, half were fraudulent and the other half were legitimate. Moreover, 16 emails were allocated to the no time pressure condition and 16 emails to the time pressure condition, with both conditions containing an equal amount of fraudulent and legitimate emails. The order in which the two conditions were shown were randomized per participant. 14 participants completed the time pressure set first (8 seconds per email) and 11 participants completed the no time pressure set first (20 seconds per email).

The outcome measures from the email legitimacy task were; the performance score on the email legitimacy task (total number of correct answers) and a confidence score (reflects extreme vs. modest scale responses).

In addition, the eye-tracking measures during the email legitimacy task acted both as predictor, as well as outcome variables. First, it was investigated whether people allocate their attentional resources differently depending on when they were or were not under time pressure. Secondly, it was desired to know whether attentional resources to specific phishing indicators within the email stimuli predict performance on the email legitimacy task. The eye-

tracking metrics, which acted as measures of attentional resources, used were as follows: visit duration (total viewing time within a specific AOI), fixation count (the number of times the gaze fixated on a specific AOI), fixation duration (how long in total the gaze was fixated on a specific AOI) and mean fixation duration (the fixation duration divided by the fixation count).

Upon completing the email legitimacy task, the participants had to perform three additional cognitive measures. These cognitive measures assessed: working memory capacity (Reading span task), inhibition (Stroop test) and cognitive reflection (Cognitive reflection task). The order in which these tasks were completed were randomized for each participant.

Materials & Procedures

Pre-experimental measures and tasks

Informed consent. Before the participants were able to partake in the study, they were required to give informed consent. The informed consent form provided the participants with a general overview of what was to be expected from them in the study, that their participation was completely voluntary and how their data would be processed and handled. The informed consent form can be found in Appendix A. After the participants indicated that they have read all their information on the form and that they accepted that their data would be used for academic purposes, they were able to proceed to the questionnaire.

Demographics questionnaire. The questionnaire can be broken down into two sections (Appendix B). In the first section, demographic information of the participants was established (gender, age, nationality and highest level of finished education), after which the participants would continue to the second section which was the email familiarity questionnaire.

Email familiarity questionnaire. To gauge the participants' experience with email services/management and their internet usage behaviour, a questionnaire was developed. This questionnaire was based on the email usage questionnaire of Jones (2016). All original questions of Jones (2016) were included to which additional questions were added. In the email familiarity questionnaire, the following questions were asked to the participants; daily time spent on the internet, proportion of time using email services, the typical number of emails received per day, preferred device for managing emails, how many phishing emails they receive in a typical week, whether they have ever responded to a phishing communication, perceived phishing knowledge and perceived vulnerability to online fraud. The following questions were added compared to the original questionnaire of Jones (2016): preferred device for managing emails and perceived vulnerability to phishing.

Experimental measures

Eye-tracking apparatus. Upon completing the email familiarity questionnaire, the eye-tracking equipment needed to be calibrated to the participant before the email legitimacy task. A Tobii Pro Fusion screen-based eye-tracker was used with a sample rate of 120Hz. In optimal conditions, the Tobii Pro Fusion achieves an accuracy of 0.3 degrees and a precision of 0.2 degrees root mean square for raw signals and 0.04 degrees after Savitsky-Golay filtering. The Tobii Pro Fusion was magnetically attached to the bottom bezel of the pc monitor. The pc monitor used was an AOC 27-inch 144hz monitor with a resolution of 1920 by 1080 pixels. Both eyes of the participant were tracked and the participants were positioned between 60-70 cm from the pc monitor/eye-tracker depending on what a comfortable position for them was to hold. No mounts were used, thus the participants were able to freely move their head. However, they were requested to minimize unnecessary movements as much as possible. Once the participants were positioned behind the pc monitor at the appropriate distance, the Tobii Pro Fusion was calibrated through the Tobii Pro Eye Tracker Manager in order to establish whether the participants were correctly seated and whether the eye-tracker was able to effectively calibrate to their eyes. Before the start of the email legitimacy task, the eye-tracker was again calibrated within the software Tobii Pro Lab using a 9-point calibration and a 4-point validation. The email stimuli were presented and data acquisition was done through Tobii Pro Lab.

Email legitimacy task. In total, the participants had to assess 32 emails for authenticity (Appendix D), of which 16 were phishing and 16 were legitimate. The 32 emails were randomly allocated in one of two sets that resulted in 16 emails per set. Both sets contained an equal amount of phishing and legitimate emails (8 legitimate and 8 fraudulent). For one set (time pressure condition), participants had 8 seconds of viewing time per email stimulus, while for the other set (no time pressure condition), participants had 20 seconds of viewing time for each email stimulus. The 8 seconds of viewing time was based on the research of Jones (2016), in which a significant negative effect of time pressure on email management decision making was found when participants had on average 8 seconds per email stimuli compared to participants that had no time limit. The 20 seconds per email stimulus was based on data that has been collected from millions of opened emails (Litmus, 2022), which stated that on average people view an email for 10 seconds. This time was doubled to ensure that participants had ample time to both read and further scrutinize the email to allow for system 2 thinking.

Email Construction. All of the emails used in this study, both legitimate and phishing, were modified genuine emails that were received by the researcher. The email stimuli were altered by replacing real with fictitious names to ensure anonymity. Only the name of the researcher, which acted as the recipient of all emails was kept in place and held constant. Additionally, the emails were selected and created to be relevant to the study sample in terms of organizations or individuals that are most likely familiar to them. Examples of this were senders such as the University of Twente, postal services (e.g., DHL) and online payment platforms (e.g., PayPal). The emails were fabricated within the Mail application of Apple which thus followed a conventional email layout. The layout design of the emails (e.g., logos, images and formats) were removed from all email stimuli, which resulted in the emails only consisting of text. This was done to limit any unwanted visual distractors that were not relevant to determine whether the emails were legitimate or phishing, since the phishing indicators used in this study were text-based. This approach was replicated from the design of McAlaney and Hills (2020).

To ensure comparability between the two sets of email stimuli, a total of 8 email templates were selected as a foundation. From these templates, 32 email stimuli were created, with 4 emails corresponding to each template. This approach ensured that the emails within a template shared similar structures and lengths. Each template consisted of 2 authentic emails and 2 phishing emails.

The phishing emails were modified to include 2 out of 4 phishing indicators: threat, urgency, suspicious links, and mismatched/generic sender domain name. The mismatched/generic sender domain name was implemented in each phishing email and combined with one of the remaining indicators. The decision was made to include the mismatched/generic sender domain name indicator in each phishing email and this choice was driven by the research objective of addressing general phishing emails rather than focusing on more sophisticated phishing schemes that involve spoofed or hacked email addresses. By including the mismatched/generic sender domain name indicator, the study aimed to capture common phishing attempts encountered by end-users in everyday email communications. Hence, this resulted in 3 different possible combinations of the phishing indicators.

For the suspicious links indicator, the phishing emails contained a fraudulent URL, while the legitimate emails contained a comparable but non-phishing indicator (i.e., a genuine URL). However, for both the urgency and threat indicators, a comparable non-phishing indicator could not be implemented.

For both the urgency and suspicious links indicators (in combination with the mismatched/generic sender domain name), 3 templates of 4 emails were made. For the threat indicator (in combination with the mismatched/generic sender domain name), 2 templates of 4 emails were fabricated. Randomization warranted that for each template 1 phishing and 1 legitimate email were allocated to a set.

AOI construction. Areas of interest (AOIs) were created for each email stimuli within the software Tobii Pro Lab. The AOIs were constructed in an identical manner for each email, hence why templates were used as described in the email construction section. Firstly, an AOI of the entire email stimulus was implemented to observe the duration of participants' examination in order to determine if the allotted time was fully used. Secondly, AOIs were made for the address bars, which included the sender details, time and date of receiving the email, subject of the email, and the details of the recipient. Thirdly, two separate AOIs were constructed for both the greeting and the signature of the email. Fourthly, for each individual paragraph (block of connected text) in the email stimulus an AOI was made. Lastly, for the phishing indicators in the fraudulent emails and the corresponding non-phishing indicators in the authentic emails two AOIs were made. The (non-)phishing indicator AOIs were located at the sender details in the address bar, which was thus a fixed location, and in the body text of which the exact location could vary.

Task procedure. Before the start of the email legitimacy task, participants were given written instructions (Appendix C) and a practice round of 2 email stimuli to familiarize themselves with the task. For one email, they had 8 seconds of viewing time and for the other they had 20 seconds. The order of the emails in the practice round were identical for each participant (8 seconds first, 20 seconds second) and eye-tracking was not used at this stage. After finishing the practice round, the participants proceeded to the email legitimacy task. First, the eye-tracker was calibrated in Tobii Pro Lab as described in the Eye-tracking apparatus section, after which the email legitimacy task would begin.

Before the start of the time pressure condition, the participants were informed that they were working under a limited amount of time, and how much time they had to analyze an email (8 seconds). Before the no time pressure condition, participants were also informed how much time they had to analyze an email stimulus (20 seconds), but it was stressed that this was ample time. For both conditions, the participants were instructed to use all the allotted time to analyze the emails before making a decision regarding the authenticity. In between the two conditions, the participants were allowed a short break and were told that they could continue the task if they felt ready to do so.

Before each email stimuli, a fixation cross in the middle of the screen was displayed that lasted exactly 1 second. After each email stimuli, the participants had to indicate on a 6-point Likert scale how confident they were that the communication was either phishing or legitimate, with -3 indicating being 'definitely phishing' and with 3 indicating being 'definitely legitimate'. The values in between indicated how confident they were in their answer. The scale was shown on screen after each email stimuli and the participants were requested to verbally provide their answer to the researcher. After providing their answer, the researcher would manually prompt the next email stimulus. It was opted to use verbal answering method to (1) limit potential answer review strategies or response pattern detection and (2) to eliminate the need for the participants to look away from the screen/eye-tracker.

Outcome measures. The outcome measures of the email legitimacy task were the total number of correct answers and a confidence score. The total number of correct answers could range between 0 and 32. In order to create a binary response, the 6-point Likert-scale was divided at the mid-point, meaning that responses between -1 to -3 were classified as phishing and responses between 1 to 3 were classified as legitimate.

The confidence score demonstrated the participant's self-confidence in their ability to accurately judge an email for its authenticity, rather than just their overall performance which was calculated in the total number of correct responses. For instance, if a participant indicated that a phishing email was 'definitely phishing' (= -3 on the Likert-scale), the confidence score for that specific email would be 3. If, on the other hand, a participant indicated that an email was 'definitely phishing' but in fact the email was authentic, their confidence score for that specific email would be -3. Total confidence scores were calculated by summing all the individual email stimulus' confidence scores and taking the average, thus resulting in a score between -3 to 3.

Post-experimental measures

Cognitive tasks. Each participant had to complete a set of three cognitive tasks after performing the email legitimacy task. During these cognitive measures, eye-tracking was not used. The motive to administer not more than three cognitive tasks was to prevent inducing fatigue effects during these tasks. This was replicated from the design of Jones (2016). The order in which the participants had to complete the cognitive tasks was randomized, to exclude potential ordering effects.

Reading Span Task. The Reading Span task assesses working memory capacity through determining how many last words of independent sentences a person can recall when

reading these sentences consecutively (Daneman & Carpenter, 1980). First, the participants had to perform 3 trials of 2 sentences and had to subsequently recall the sentences' last words. If the participant answered 2 out of the 3 trials correctly, then the number of sentences increased by 1. The test continued in this manner until the participant could not recall at least 2 out of the 3 trials correctly, or when the maximum level of 6 sentence trials was achieved. The reading span size for each participant was determined based on the highest number of sentences' last words that they were able to recall. So, if a participant was able to recall at least 2 out of 3 trials with 3 sentences, but was unable to do so for 4 sentences, their reading span size would be 3. The possible scores thus range from 2 to 6.

In total, 66 sentences were taken from Daneman and Carpenter (1980) to use as stimuli in this task (Appendix E). The sentences were displayed via PowerPoint, using the same font and letter size for each stimulus. Each participant completed the task identically, which implies that the order of the sentences was not randomized or changed. Before the start of a trial, the participants were shown a slide indicating which trial number they were on and the number of sentences the participants had to recall the last words from. After this slide, the first sentence was displayed. The participants had to read the sentence out loud, but were allowed to do so at their own pace. When the first sentence was fully read, the researcher prompted the next sentence and this process was repeated until all sentences in the trial were read. After each trial, a blank white screen was shown and the participants had to verbally provide their answer to the researcher. Before the start of the reading span task, the participants were given 3 practice trials of 2 sentences to familiarize themselves with the task.

Stroop Test. The Stroop test is used as a behavioural measure of inhibition (Stroop, 1935). With the current version of the Stroop test, participants had to match the font colour of a word with the corresponding colour that is binded to a key on a keyboard. However, the word of the colour and the colour of the font the word is written with, can either be congruent (e.g., the word is blue and the font is blue) or incongruent (e.g., the word is blue, but the font is in red). By doing this, participants were required to suppress the word stimuli in order to provide the correct response, namely the colour of the font to the correct colour key binding.

The Stroop test was administered through PsyToolkit (Stoet, 2010; Stoet, 2017). In total, there were 9 different word and colour combinations possible. The word and font colours were red, green and blue, with the corresponding key bindings of r, g and b on a keyboard. It was opted to use these key bindings, since they (1) each represent the first letter of a used colour to avoid confusion and (2) are positioned near each other on the keyboard to prevent participants needing to find the correct key while working on the test (Jones, 2016).

The word stimuli were displayed on a black background and a fixation cross was shown before a stimulus was presented. Between the word stimuli, a delay of 500 milliseconds was built-in. If a participant took longer than 2000 milliseconds to respond to a certain stimulus, the program would automatically define it as a wrong answer and would continue to the next stimulus. Replicating the design of Jones (2016), the participants first completed 27 practice trials to familiarize themselves with the test and the inputs on the keyboard. After the practice trials, the participants completed 144 trials that were split in two sections of 72 trials, with a short break in between (Jones, 2016). The scores on the Stroop test were determined by taking the mean difference between the response times on the incongruent and congruent trials that had correct responses. A bigger mean difference implicates that the participant required more time to differentiate between congruent and incongruent stimuli, hence signifying lower levels of inhibition.

Cognitive Reflection task. The Cognitive Reflection Task is a relatively brief test which involves solving three problems (Appendix F) (Frederick, 2005). The principle of this test, is that each problem has an intuitive response, which is not the correct answer. The test was originally developed for an American audience, since dollars were used as currency. Hence, this was replaced by euros to match the setting of the current study. An example question of the Cognitive Reflection Task is: "A bat and a ball cost €1.10 in total. The bat costs €1.00 more than the ball. How much does the ball cost?". The intuitive answer to this question would be 10 cents, while the correct answer is 5 cents. The test score of the participants was the number of correct responses to the question, which thus ranges from 0 to 3. There was no time limit set for the participants to answer the questions, but they were informed that the task usually does not take longer than 3 minutes to complete. The questions were administered to the participants through Qualtrics.

Results

Email legitimacy task performance and time effects

Descriptives

The email stimuli were first examined for variations in difficulty. Appendix G presents the means of the 6-point Likert scale responses for each email stimulus. The results indicate that participants faced varying levels of difficulty in correctly classifying the emails. Out of the 32 emails, 5 were deemed as rather difficult, while 8 were considered relatively easy.

Difficult emails were consistently misjudged by participants which placed them on the wrong side of the midpoint on the scale, while easy emails had a mean scale response equal or larger than -2 or +2 depending on the type of email. In addition, the participants' responses were analyzed for bias in labeling email stimuli as phishing or legitimate. Since there were an equal number of phishing and legitimate emails in the task, it is expected that someone who does not exhibit a response bias would have a mean scale response between -1 and 1 (on the 6-point Likert scale between -3 and 3). Furthermore, a large standard deviation of the scale responses would indicate that the participant has used the entire scale when rating the emails on authenticity, which thus suggests a lower response bias. Based on these conditions, no bias was found, as mean scale responses for all emails combined were not greater than -1 or 1 and standard deviations were (apart from two cases) above 1.5.

Table 1 shows means and standard deviations for outcome measures of the email legitimacy task, while Table 2 shows Pearson r correlations between these measures. What can be observed is that the total confidence score strongly correlates with performance on the email legitimacy task, while the confidence scores on phishing and authentic emails separately (although less strong) are also strongly correlated with email decision-making performance. Notably, the outcome measures of phishing and authentic emails are not correlated, indicating that performance and confidence for determining if an email is legitimate does not correlate with performance and confidence in detecting if an email is a phishing email, and vice versa.

Table 1.

Descriptive statistics for measures of the email legitimacy task

Measure	M	SD
Total correct (/32)	22.76	2.83
Authentic correct (/16)	11.60	2.06
Phishing correct (/16)	11.20	2.35
Total confidence score (-3 to 3)	1.01	0.46
Confidence score authentic (-3 to 3)	1.02	0.58
Confidence score phishing (-3 to 3)	0.99	0.69

Table 2.*Pearson correlations between the outcome measures of the email legitimacy task*

	1	2	3	4	5	6
1. Total correct	1.00					
2. Authentic correct	.60**	1.00				
3. Phishing correct	.70**	-.15	1.00			
4. Total confidence score	.95**	.59**	.66**	1.00		
5. Confidence score authentic	.63**	.97**	-.073	.67**	1.00	
6. Confidence score phishing	.76**	-.013	.95**	.78**	.068	1.00

***. Correlation is significant at $p < .01$ level (2-tailed).*

To determine whether the participants performed better on the email legitimacy task than would be expected by chance, one-sample t-tests were performed. The one-sample t-tests indicated that the participants performed better on the entire email legitimacy task than would be expected by chance ($t(24) = 11.93, p < .001, d = 2.39$), as well as for phishing emails ($t(24) = 6.82, p < .001, d = 1.36$) and authentic emails ($t(24) = 8.73, p < .001, d = 1.75$) separately.

Inferential statistics – Hypothesis 1

To assess the effect of time pressure on email management decision-making, paired sample t-tests were performed. These indicate that participants during the time pressure condition ($M = 11.60, SD = 1.61$) were not less accurate in correctly classifying emails when compared to the no time pressure condition ($M = 11.16, SD = 1.99$), $t(24) = 0.98, p = 0.339, d = 0.20$. Moreover, a significant difference was not found when assessing the differences in confidence scores between the time pressure ($M = 1.03, SD = 0.45$) and no time pressure ($M = 0.98, SD = 0.60$) conditions, $t(24) = 0.50, p = 0.624, d = 0.10$. Hence, these results do not support the hypothesis that inducing time pressure lowers performance on email management decision-making as measured by the email legitimacy task.

To determine whether the order of the time conditions in which the participants completed the email legitimacy task had an effect on performance, an independent t-test was performed. The outcome of this test suggests that either performing the time pressure condition ($N = 14, M = 22.71, SD = 2.56$) or the no time pressure condition ($N = 11, M =$

22.82, $SD = 3.28$) first had no significant effect on the performance on the email legitimacy task, $t(23) = -0.09$, $p = 0.930$, $d = -0.04$.

It was also examined through multiple regression analysis whether responses to the email familiarity questionnaire predicted performance on the email legitimacy task. For brevity, Table H1 in Appendix H provides the descriptive statistics to each question in the questionnaire. None of the questions were found to be significant predictors of accuracy. The model that included all questions had a negative adjusted R^2 of -0.02 ($F(8, 16) = 0.94$, $p = 0.515$), while a model which only included the questions regarding perceived ability to detect phishing emails and perceived risk to online fraud produced an adjusted R^2 of 0.08 ($F(2, 22) = 2.05$, $p = 0.153$). The regression tables can be found in Table H2 and H3 in Appendix H.

Cognitive measures

Descriptives

The distributions of the outcome measures of the three cognitive measures were first examined for floor/ceiling effects. The distribution of reading span scores showed positive skewness, with the lowest possible score ($= 2$) having the highest frequency, indicating the presence of a floor effect. The distributions of the remaining two cognitive measures did not show clear evidence of such an effect. Table 3 shows the descriptive statistics of the cognitive measures. Note that five of the participants indicated that they knew the answer to one or more of the questions in the cognitive reflection test, so their scores were excluded from the analyses. However, their scores on the other cognitive measures were included. Table 4 shows the Pearson r correlations between cognitive measures and outcome measures of the email legitimacy task, suggesting no significant relationship between them. Moreover, the cognitive tasks were also not correlated to each other.

Inferential statistics – Hypotheses 2-4

Multiple regression analyses were conducted to test whether the cognitive measures predicted outcome measures of the email legitimacy task. Separate analyses were run for each outcome measure of the email legitimacy task. However, none of the cognitive measures significantly predicted the outcome measures of the email legitimacy task, even when controlling for the different time conditions. Results are summarized in Table 5. Hence, the hypotheses, which stated that people with a (H2) lower working memory capacity, (H3) lower inhibitory capacity and (H4) lower cognitive reflection would perform with lower performance on the email legitimacy task, cannot be supported.

Table 3.*Descriptive statistics for cognitive measures*

	<i>N</i>	<i>M</i>	<i>SD</i>
Cognitive reflection test (0-3)	20	1.15	1.09
Reading span task (2-6)	25	2.88	0.88
Stroop effect (in milliseconds)	25	47.32	48.67

Table 4.*Pearson correlations between the outcome measures of the email legitimacy task and the cognitive measures*

	1	2	3	4	5	6	7	8	9
1. Total correct	1.00								
2. Total confidence score	.95**	1.00							
3. Correct time pressure	.73**	.71**	1.00						
4. Correct no time pressure	.83**	.78**	.23	1.00					
5. Confidence score time pressure	.81**	.85**	.92**	<i>.41*</i>	1.00				
6. Confidence score no time pressure	.87**	.92**	<i>.41*</i>	.91**	.56**	1.00			
7. Cognitive reflection	.39	.42	.30	.37	.36	.40	1.00		
8. Reading span	.22	.23	.05	.27	.05	.32	.44	1.00	
9. Stroop score	-.30	-.21	-.04	-.39	-.14	-.21	-.34	-.14	1.00

***. Correlation is significant at the $p < .01$ level (2-tailed).*

**. Correlation is significant at the $p < .05$ level (2-tailed).*

Table 5.

Multiple regression analyses of the cognitive measures predicting the outcome measures of the email legitimacy task split by time pressure conditions

Model		<i>B</i>	<i>SE</i>	β	<i>t</i>	<i>p</i>	Adj. <i>R</i> ²
Total	(Constant)	21.81	2.37		9.21	<.001	.04
Correct	Cognitive reflection	0.77	0.69	0.30	1.13	.277	
	Reading span	0.20	0.81	0.06	0.25	.804	
	Stroop score	-0.01	0.01	-0.19	-0.79	.438	
Total	(Constant)	0.77	0.39		1.96	.067	.03
Confidence score	Cognitive reflection	0.15	0.11	0.36	1.37	.191	
	Reading span	0.03	0.13	0.06	0.25	.806	
	Stroop score	-0.001	0.002	-0.07	-0.31	.763	
Correct	(Constant)	11.39	1.41		8.07	<.001	-.07
Time pressure	Cognitive reflection	0.54	0.41	0.37	1.33	.204	
	Reading span	-0.18	0.48	-0.01	-0.38	.712	
	Stroop score	0.002	0.008	0.06	0.27	.795	
Correct	(Constant)	10.41	1.62		6.42	<.001	.09
No time pressure	Cognitive reflection	0.37	0.47	0.20	0.78	.448	
	Reading span	0.32	0.55	0.14	0.58	.573	
	Stroop score	-0.01	0.01	-0.30	-1.30	.212	
Confidence score	(Constant)	1.04	0.39		2.68	.016	-.02
Time pressure	Cognitive reflection	0.17	0.11	0.41	1.52	.148	
	Reading span	-0.07	0.13	-0.13	-0.50	.621	
	Stroop score	<0.001	0.002	-0.02	-0.009	.926	
Confidence score	(Constant)	0.49	0.50		0.99	.336	.04
No time pressure	Cognitive Reflection	0.16	0.14	0.29	1.09	.294	
	Reading span	0.12	0.17	0.18	0.73	.474	
	Stroop score	-0.001	0.002	-0.09	-0.35	.728	

Eye-tracking measures

Experimental manipulation check

The mean visit duration of the total image AOIs was calculated for all email stimuli to determine if participants used the full allotted time to analyze the emails in both time conditions. One participant was excluded from the analyses due to missing, on average, 25% of eye-tracking data in the time pressure condition. In order to compare the time pressure condition ($M = 7.86$, $SD = 0.26$) with the no time pressure condition ($M = 19.73$, $SD = 0.59$), proportions were calculated by dividing the total time participants spent looking at email stimuli of each condition by the total time available. Results showed no difference between the time pressure ($M = 0.98$, $SD = 0.03$) and the no time pressure ($M = 0.99$, $SD = 0.03$) conditions in terms of proportion of time viewing email stimuli, $t(23) = -0.74$, $p = 0.469$, $d = -0.15$.

To determine whether the time manipulation had an impact on how long participants fixated on the phishing indicators, paired t-tests on the fixation duration (in seconds) of the four phishing indicators per time condition of the fraudulent emails. For brevity, the results of these t-tests are presented in Appendix I. The findings indicated that participants in the no time pressure condition, on average, fixated significantly longer on each of the phishing indicators compared to those in the time pressure condition. Therefore, this suggests that the time manipulation led to significant differences in fixation duration regarding the relevant regions of the fraudulent emails, confirming that the time pressure manipulation did effectively manipulate the attention allocated to each phishing indicator prior to participants making a judgement.

Attentional resources to the phishing indicators – Hypothesis 5

The purpose of this section is to investigate whether individuals allocate their attentional resources differently to phishing indicators when faced with time constraints versus having more time. Attentional resource allocation to phishing indicators in this study was defined as how long in total participants fixated their gaze on a specific phishing indicator AOI (fixation duration), how often their gaze fixated to the specific phishing indicator AOI (fixation count) and how long the average fixation to the phishing indicator AOI lasted (mean fixation duration). To control for the time difference between the conditions, proportions of the fixation duration and the fixation count variables of the no time pressure condition were computed by dividing the values by 2.5 (proportional difference

between the time conditions). The mean fixation duration was acquired by dividing the fixation duration by the fixation count.

Firstly, paired sample t-tests were performed on the fixation duration of time pressure variables and the proportionalized no time pressure variables for each phishing indicator AOI. The results, presented in Table 6, indicate that participants did not spend significantly more time fixating on the mismatched/generic sender domain name and suspicious link AOIs in either time condition. However, for the urgency and threat AOIs, participants proportionally fixated longer in the time pressure condition than in the no time pressure condition.

Table 6.

Paired sample t-tests of fixation duration on the phishing indicator AOIs controlled for time.

		<i>M</i>	<i>SD</i>	<i>t</i>	<i>df</i>	<i>p</i>
Mismatched/Generic sender domain name	Time pressure	1.20	0.79	-0.81	23	.425
	No time pressure	1.27	0.68			
Suspicious links	Time pressure	0.38	0.36	0.17	23	.864
	No time pressure	0.36	0.28			
Urgency	Time pressure	0.63	0.32	2.15	23	.042*
	No time pressure	0.49	0.20			
Threat	Time pressure	0.57	0.36	2.18	23	.040*
	No time pressure	0.41	0.16			

*. *Significant at the $p < .05$ level (2-tailed).*

Secondly, differences in fixation count were examined. Paired sample t-tests were again conducted. The results, presented in Table 7, indicate that participants did not significantly differ in the number of fixations on the mismatched/generic sender domain name and suspicious links AOIs in both conditions when controlling for time differences. However,

the participants fixated more frequently on the urgency and threat phishing indicators in the time pressure condition compared to the no time pressure condition. This is the same pattern that was observed for the fixation duration.

Table 7.

Paired sample t-tests of fixation count on the phishing indicator AOIs controlled for time.

		<i>M</i>	<i>SD</i>	<i>t</i>	<i>df</i>	<i>p</i>
Mismatched/Generic sender domain name	Time pressure	5.61	3.54	0.29	23	.771
	No time pressure	5.47	2.89			
Suspicious links	Time pressure	1.97	1.66	0.35	23	.731
	No time pressure	1.83	1.19			
Urgency	Time pressure	3.83	1.95	2.72	23	.012*
	No time pressure	2.80	1.09			
Threat	Time pressure	3.50	2.27	2.54	23	.018*
	No time pressure	2.37	0.77			

*. *Significant at the $p < .05$ level (2-tailed).*

Thirdly and lastly, it was examined whether there was a difference in time participants on average fixated to the phishing indicator AOIs between the time conditions. The mean fixation duration of both time conditions was compared by subjecting them to a series of paired sample t-tests. The results of these t-tests can be found in Table 8. As can be observed from the results, a significant difference was found in mean fixation duration on mismatched/generic sender domain name AOIs, where in the no time pressure condition the mean duration of the fixations was higher compared to the time pressure condition. For the remaining three phishing indicators, no significant differences were found between the time conditions.

Table 8.*Paired sample t-tests of mean fixation duration on the phishing indicator AOIs.*

		<i>M</i>	<i>SD</i>	<i>t</i>	<i>df</i>	<i>p</i>
Mismatched/Generic sender domain name	Time pressure	0.21	0.03	-3.88	23	<.001**
	No time pressure	0.23	0.03			
Suspicious links	Time pressure	0.16	0.07	-0.76	23	.458
	No time pressure	0.17	0.06			
Urgency	Time pressure	0.17	0.02	-1.68	23	.107
	No time pressure	0.18	0.03			
Threat	Time pressure	0.19	0.24	0.04	23	.968
	No time pressure	0.19	0.18			

***. Significant at the $p < .01$ level (2-tailed).*

**. Significant at the $p < .05$ level (2-tailed).*

Together these results indicate that when under time pressure participants spent more of the available time looking at phishing indicators of threat and urgency (fixation duration), and looked at these indicators more often (fixation count). Being under time pressure did not change how long and how often people fixated looking at any of the other phishing indicators. However, when the participants were not under time pressure, the average duration of each separate fixation to the mismatched/generic sender domain name phishing indicator lasted significantly longer compared to when being under time pressure (mean fixation duration). Therefore, hypothesis 5, which stated that participants in the time pressure condition would provide less attentional resources to the phishing indicators, cannot be supported with the current results. Furthermore, across all eye-tracking metrics, participants consistently devoted

the most attentional resources to the mismatched/generic sender domain name phishing indicator and the least to the suspicious links phishing indicator.

Eye-tracking metrics on email legitimacy task performance – Hypothesis 6

To determine whether the outcome measures of the email legitimacy task were associated with the eye-tracking metrics of the phishing indicator AOIs, Pearson r correlations were calculated and can be found in Table 9. The eye-tracking metrics variables were grouped by their corresponding phishing indicator AOIs. To group the variables, the mean of the time pressure variable and the proportionalized no time pressure variable was taken.

As can be observed from Table 9, the amount of time and number of fixations on the mismatched/generic sender domain name phishing indicator AOIs were significantly positively correlated with both the total correct answers and total confidence score of the email legitimacy task. However, when considering only the correct responses and confidence scores to the phishing email stimuli, these correlations were not significant. The duration and the number of fixations on the suspicious links phishing indicator AOIs did not correlate with any of the outcome measures of the email legitimacy task. In contrast, the urgency and threat phishing indicator AOIs were significantly negatively correlated with the total correct answers and total confidence score, indicating that spending more time and fixating more often on these indicators is associated with lower performance and confidence in judgement.

These findings partly support hypothesis 6, which proposed that the attentional resources given to phishing indicators are associated with judgement performance on the email legitimacy task. How long and frequently people looked was associated with performance on the email legitimacy task for three out of the four phishing indicators. The mismatched/generic sender domain name indicator helped people to successfully classify emails, while the urgency and threat cues made people less accurate on the task. The suspicious links indicator had no clear effect on performance. By combining these observations, it could be inferred that when participants spent less time and frequency fixating on the mismatched/generic sender domain name phishing indicator, and instead focused more on the threat and urgency indicators (or other parts of the email), their performance and confidence in making judgments may decrease.

Table 9.

Pearson correlations for the eye-tracking methods on the phishing indicators AOIs on outcome measures of the email legitimacy task

	1	2	3	4	5	6	7	8	9	10	11	12
1. Total correct	1.00											
2. Phishing correct	.70**	1.00										
3. Total confidence score	.95**	.66**	1.00									
4. Confidence score phishing	.76**	.95**	.78**	1.00								
5. Fixation duration Mis/Gen sender	.52**	.25	.49*	.29	1.00							
6. Fixation duration Suspicious links	-.01	-.38	.09	-.21	-.13	1.00						
7. Fixation duration Urgency	-.44*	-.30	-.51*	-.41*	-.45*	.10	1.00					
8. Fixation duration Threat	-.53**	-.42*	-.57**	-.51*	-.52**	.06	.76**	1.00				
9. Fixation count Mis/Gen sender	.54**	.28	.52**	.31	.98**	-.14	-.41*	-.53**	1.00			
10. Fixation count Suspicious links	-.04	-.34	.08	-.17	-.17	.98**	.13	.08	-.15	1.00		
11. Fixation count Urgency	-.41*	-.26	-.44*	-.36	-.41*	.11	.97**	.71**	-.34	.16	1.00	
12. Fixation count Threat	-.45*	-.32	-.47*	-.41*	-.50*	.08	.74**	.95**	-.47*	.15	.74**	1.00

****.** Correlation is significant at the $p < .01$ level (2-tailed).

***** Correlation is significant at the $p < .05$ level (2-tailed).

Discussion

The aim of this study was to examine whether situational and individual factors, that have been related to phishing susceptibility in the scientific literature, had an impact on email management decision-making. In addition, the study used eye-tracking technology to capture viewing behaviour in order to observe whether differences in attentional resource allocation to phishing indicators predicted email judgement performance, and whether experimental manipulations affected attentional resource allocation. The results revealed that inducing time pressure did not lead to worse performance on the email legitimacy task, and the hypothesized relationships between cognitive factors (working memory capacity, inhibition, and cognitive reflection) and phishing susceptibility were not supported. Analysis of the eye-tracking data demonstrated that, after controlling for time differences, participants under time pressure exhibited longer and more frequent fixations on threat and urgency phishing indicators, in comparison to the no time pressure condition. Conversely, participants under no time pressure displayed longer mean fixation durations on the mismatched/generic sender domain name phishing indicator, as opposed to being under time pressure. Allocating more attentional

resources to threat and urgency phishing indicators was negatively associated to judgement performance and confidence, whereas a positive association was observed for the mismatched/generic sender domain name phishing indicator. The suspicious links phishing indicator did not exhibit a clear effect on email judgement performance or confidence and consistently received the least attentional resources out of all phishing indicators regardless of time. In contrast, the mismatched/generic sender domain name phishing indicator consistently attracted the most attentional resources across both time conditions.

Main findings

Time conditions on email judgement performance

The present study did not find evidence to support the notion that imposing time constraints negatively affects decision-making performance in an email legitimacy task. Previous research often links individual differences in phishing susceptibility to dual process theories, suggesting that relying on intuitive thinking increases the likelihood of making inaccurate judgements about email authenticity (Dong et al., 2008; Jones et al., 2015; Vishwanath et al., 2011). To encourage intuitive thinking, researchers commonly impose time constraints on participants during email management tasks (Yan & Gozu, 2012; Jones, 2016).

The specific implementation of time constraints in this study differed from previous research. In this study, participants were given a fixed time limit per email stimulus, while other studies imposed time constraints over the entire task (Yan & Gozu, 2012; Jones, 2016). It was opted to use a fixed time limit per email stimulus in both time conditions in order to be able to make valid comparisons between and within time conditions regarding the eye-tracking metrics. Nonetheless, this discrepancy in defining time conditions may have impacted the results.

Literature suggests that by just perceiving time pressure, it can weaken rational thinking and lead to more intuitive decisions (Dijker & Koomen, 1996; Finucane et al., 2000). However, the participants in this study did not significantly differ in their confidence scores between the time conditions. This could imply that participants did not feel more or less pressured in either time condition, and therefore did not have to rely on an intuitive mode of thinking when making judgements. As a result, the time constraints did not have a significant effect on task performance. However, additional research is required to investigate the influence of diverse techniques of imposing time constraints on email management tasks and how these constraints ultimately impact judgement performance.

Cognitive factors on email judgement performance

The influence of cognitive factors, specifically working memory capacity, inhibition, and cognitive reflection, on judgement performance in the email legitimacy task was assessed. In addition, it was examined whether the effects of these cognitive factors on email judgement performance were influenced by the presence of time constraints. Contrary to initial expectations, the results did not provide evidence to support the impact of these cognitive factors on email judgement performance, regardless of whether participants worked under time constraints or not. These findings deviate from previous research that has linked these cognitive factors to phishing susceptibility and dual process theories (Mayhorn & Nyeste, 2012; Jones et al., 2015; Jones, 2016). It has been suggested in the scientific literature that individuals with lower abilities in these factors are more likely to rely on an intuitive mode of thinking, thus resulting in a higher probability of making judgement errors during email management decision-making.

There are several possible explanations as to why the cognitive measures were not related to judgement performance in the current study. Firstly, while the Stroop test and reading span task were administered as similarly as possible to Jones' (2016) study, the tasks were administered on different platforms, which may have caused slight variations in these specific tasks. However, it is unlikely that these slight variations had a large impact on the results, since the core task procedures were not changed. Secondly, this study's participants were non-native English speakers. Although language proficiency was likely not a limiting factor for the Stroop test and cognitive reflection task, research suggests that it could have a negative impact on verbal working memory capacity tasks like the reading span task (Linck et al., 2014). However, in this study, participants performed better on average on the reading span task than those in Jones' (2016) study, which casts doubt on this explanation. Lastly, it is possible that the single task of judging emails for authenticity in this study was not sufficiently cognitively demanding enough for the cognitive factors to have an effect on performance. Furthermore, the non-significant differences in confidence scores between the time conditions suggest that participants did not experience differences in pressure in either time condition. This could explain the lack of observable effects of the cognitive tasks under time constraints, as these were hypothesized to be connected to the dual process theories. Therefore, at this stage, no definite conclusions can be drawn about whether these cognitive factors have an impact on email management decision-making and how these effects might be influenced by situational factors such as time pressure.

Attentional resource allocation to phishing indicators

Threat and urgency. As discussed, there was no difference observed in email judgement performance between the time conditions. However, there were significant differences in attentional resource allocation to phishing indicators between the time conditions. It was found that when participants were under time pressure, they would proportionally fixate longer and more frequent to threat and urgency indicators compared to the no time pressure condition.

Previous eye-tracking research has shown that when individuals have limited time to read information, they often adopt a skimming strategy to cope with the volume of text (Duggan & Payne, 2009; Duggan & Payne, 2011). It was found that this skimming approach directs people's visual attention towards the most important information in the text. In the study by McAlaney and Hills (2020), participants were found to prioritize phishing indicators of threat and urgency over other indicators, such as misspellings. This suggests that indicators related to threat and urgency may evoke a survival information bias, which proposes that individuals prioritize information relevant to their well-being (Nairne, 2010). Therefore, it is possible that participants, when constrained by time, adopt a different reading strategy that involves skimming through the email, while proportionately allocating more attentional resources to phishing indicators of threat and urgency that relate to their well-being.

On the other hand, phishing indicators of threat and urgency were also negatively associated with email judgement performance, meaning that the longer and more often participants fixated their gaze to these indicators, email judgement performance decreased. Previous research regarding phishing susceptibility has revealed that urgency indicators make people more likely to respond to phishing emails than those without (Cui et al., 2020). The explanation for this is that urgency indicators are known to use large amounts of information processing resources (Shah et al., 2004). Moreover, individuals who disproportionately focus on these urgency indicators, while overlooking other elements of the email, are at a higher risk of falling victim to phishing attacks (Marett & Wright, 2009; Vishwanath et al., 2011; Cui et al., 2020). When combining the results of this study with the available scientific literature, it appears likely that pressuring techniques in phishing emails such as urgency and threat indicators require a large amount of information processing resources. Consequently, when participants disproportionately focused their attention on these indicators, they allocated less attention to other elements of the email, thus increasing their likelihood to misjudgment.

Mismatched/generic sender domain name. In contrast to the threat and urgency phishing indicators, the results revealed a positive association between eye-tracking metrics

and judgment performance on the email legitimacy task for the mismatched/generic sender domain name phishing indicator. This suggests that the sender information played an important role in correctly classifying emails. Previous research provides support for this notion, which indicated that end-users highly prioritize the sender information when judging emails (Downs et al., 2006). The study by Downs et al. (2006) indicated that 95% of participants reported using the "from" field to identify any discrepancies between the sender's address and the sender's name. Furthermore, it was found that participants consistently allocated most attentional resources to the mismatched/generic sender domain name indicator compared to the other three phishing indicators. This finding aligns with previous eye-tracking research regarding phishing emails, where participants tended to focus most on the sender information if this is visually available when making veracity judgements (Pfeffel et al., 2019).

Prior research has also demonstrated that people's ability to correctly identify phishing emails can be improved through the use of nudges, which aim to direct individuals' gaze towards the sender information using visual aids (Nicholson et al., 2017; Huang et al., 2022). However, it is important to note that even among participants who already allocated the most attentional resources to the sender information in this study, none were able to correctly identify all phishing emails. This emphasizes the importance of not only directing attention to relevant aspects of an email, but also providing individuals with the necessary knowledge and strategies to effectively recognize phishing indicators to consistently make correct judgments. Xiong et al. (2017) demonstrated this in their study, showing that highlighting domain names in web page URLs did not offer effective protection against phishing, indicating a lack of end-users' knowledge of web page security indicators or how to use them effectively.

Additionally, it is worth noting that the sender information is often hidden by default in popular email clients. For instance, in Google's Gmail web interface, end-users need to hover over the sender's name to be able to view the sender information. Similarly, in Microsoft Outlook, multiple steps need to be performed in order to access the sender information (Nicholson et al., 2017). These practices are unlikely to assist end-users in correctly identifying phishing attacks and should therefore be avoided.

Suspicious links. The suspicious links phishing indicator consistently received the least amount of attentional resources from the participants compared to other phishing indicators. Moreover, this indicator was not significantly associated with email judgment performance. Prior research has highlighted the importance of examining URLs when assessing the legitimacy of emails, as individuals who scrutinize URLs are less likely to click

on phishing links (Downs et al., 2007). Anti-phishing trainings have also emphasized the significance of scrutinizing embedded URLs in emails to reduce the risk of falling victim to phishing attacks (Kumaraguru et al., 2007; Kumaraguru et al., 2010; Sheng et al., 2007). The discrepancy between the literature's emphasis on URL examination and the lack of visual attention and association to judgement performance found in this study raises questions about the reasons behind this observation.

There is currently no research specifically examining the interaction between (fraudulent) URLs within emails and email management decision-making. However, one study using eye-tracking technology investigated how people visually process (fraudulent) URLs and found that the presence of "www" in the domain name was perceived as a safety indicator, resulting in less attention being given to the rest of the URL (Ramkumar et al., 2020). Additionally, it was observed that people have a cognitive resource limit, typically around 100 characters, beyond which additional time is not allocated for examining the URL. Although the study focused solely on URL analysis, it is highly likely that various URL characteristics influence how people visually attend to them (Ramkumar et al., 2020). In the present study, the URLs used were under 100 characters, but all included "www" in the domain name. This visual characteristic of the URLs may have led participants to underestimate the importance of this indicator in detecting phishing attempts, as indicated by the limited attentional resources allocated. However, further research is needed to explore how the presence and appearance of URLs in (phishing) emails impact email management decision-making, to determine the reliability of the findings in this study.

Practical implications

The findings in this study have important practical implications for both technological and human interventions. On the technological front, interventions should prioritize adjustments to email clients to prominently display sender information. This study demonstrates that when sender information is visually accessible to end-users, it receives the most attentional resources during email management decision-making, which is related to improved judgement performance. Previous research supports this notion, but also highlights the effectiveness of nudges in the form of visual aids to guide end-users' attention towards relevant aspects of an email, such as sender information, thereby enhancing judgement performance (Nicholson et al., 2017; Huang et al., 2022). However, it is important to consider potential desensitization effects of these nudges over time due to repeated exposure, which may diminish their effectiveness (Vitek & Syed Shah, 2019; Shah et al., 2021). Therefore,

careful implementation of nudges in email systems is crucial to maximize their benefits. Furthermore, this study revealed that URLs in emails receive the least attention when making email legitimacy judgments, despite prior research indicating that examining URLs is a reliable indicator for identifying phishing attempts (Downs et al., 2007). Similar to the sender information, nudging end-users to pay attention to (embedded) URLs within emails could potentially enhance judgement performance.

However, the effectiveness of these technological measures is likely contingent upon the phishing knowledge of end-users. Simply highlighting or prominently displaying relevant email components indicative of phishing may have limited utility if end-users are unable to interpret them correctly (Xiong et al., 2017). Therefore, human-centered interventions should continue to prioritize educating individuals about phishing risks and how to identify phishing attempts, as various studies have shown that such education can effectively reduce victimization rates (Sheng et al., 2007; Kumaraguru et al., 2010; Dodge et al., 2012).

Limitations

There were a number of limitations to this study. A relatively small sample size was used (N= 25). However, this is not uncommon when compared to other eye-tracking studies regarding phishing email susceptibility (Pfeffel et al., 2019; McAlaney & Hills, 2020). While the sample size is not necessarily smaller than found in other comparable eye-tracking studies, it should be noted that this sample size makes it difficult to identify anything other than very large effects. Moreover, the participants also consisted of a narrow demographic, which were all students at the University of Twente with an age range of 18-26 years old. Gender in the sample was almost equally distributed with 56% female and 44% male. Although, from scientific literature there is no consistent evidence that both gender and age have an impact on phishing susceptibility (Jagatic et al., 2007; Sheng et al., 2010; Kumaraguru et al., 2010; Darwish et al., 2012; Halevi et al., 2015; Liu et al., 2020, Mohebzada et al., 2012). Moreover, there is no evidence to suggest gender differences in eye movements (Klein and Ettinger, 2019).

The email stimuli used in this study did not resemble the visual appearance of real-life phishing emails. It is important to acknowledge that visual elements can significantly affect email management decision-making. For instance, research has shown that when phishing is designed to look like legitimate messages from a trusted source, end-users are more likely to be deceived by them (Egelman et al., 2008). Similarly, Kumaraguru et al. (2007) found that end-users were more likely to click on links in phishing emails that contained logos and

images of trusted sources. However, for this study, it was opted to remove all visual elements from the email stimuli to keep them as comparable as possible to each other and to reduce any unnecessary visual distractors. This research only used text-based phishing indicators, hence, removing all other visual elements would ensure the highest possible internal validity. Nonetheless, it is important to note that the absence of visual elements in the email stimuli may not fully reflect the reality of phishing attacks.

Finally, it should be noted that instructing the participants to perform an email legitimacy task could have influenced the results. Providing participants in advance information about the task they are about to engage in may potentially introduce a bias in the results, as it could prompt them to adopt a more deliberate and rational thinking process when evaluating the emails (Parsons et al., 2015; Jones et al., 2015). This heightened awareness and cognitive processing might not accurately represent the spontaneous and automatic nature of real-life scenarios. Thus, it is important to acknowledge that pre-informing participants about the task could impact the ecological validity of the study, potentially limiting the generalizability of the findings to real-world contexts.

Conclusion

This study highlights the effectiveness of eye-tracking technology in examining the relationship between visual attention to phishing indicators and email management decision-making performance. The findings suggest that providing more visual attention to sender information is positively associated with email judgement performance, while a negative association was observed for threat and urgency indicators. However, visual attention to suspicious URLs within emails did not show a clear effect on performance. These findings emphasize the importance of prominently displaying sender information in email clients, potentially supplemented with visual aids to guide end-users' attention towards this aspect. Additionally, this study revealed the hazards social engineering tactics (of threat and urgency) can pose. These tactics are inherently designed to attract attention and it was demonstrated that the more visual attention they received related to worse email judgement performance. This effect may be particularly relevant when individuals are under time constraints, as indicated by the proportionally higher visual attention allocated to threat and urgency indicators in these conditions. Building upon this research can provide more valuable insights into visual attention to phishing indicators in relation to email judgement performance, and factors influencing this process. This knowledge could contribute to the development of more effective technological solutions and interventions for educating end-users.

References

- 2021 state of email engagement. Litmus. (2022, August 16). Retrieved December 14, 2022, from <https://www.litmus.com/resources/state-of-email-engagement/>
- Abu-Nimeh, S., Nappa, D., Wang, X., & Nair, S. (2007). A comparison of machine learning techniques for phishing detection. In Proceedings of the anti-phishing working groups 2nd annual eCrime researchers summit (pp. 60-69).
- Almomani, A., Gupta, B. B., Atawneh, S., Meulenberg, A., & Almomani, E. (2013). A survey of phishing email filtering techniques. *IEEE communications surveys & tutorials*, 15(4), 2070-2090.
- Atkins, B., & Huang, W. (2013). A study of social engineering in online frauds. *Open Journal of Social Sciences*, 1(03), 23.
- Bergholz, A., De Beer, J., Glahn, S., Moens, M. F., Paaß, G., & Strobel, S. (2010). New filtering approaches for phishing email. *Journal of computer security*, 18(1), 7-35.
- Bullée, J. W., & Junger, M. (2020). Social engineering. *The Palgrave Handbook of International Cybercrime and Cyberdeviance*, 849-875.
- Button, M., Lewis, C., & Tapley, J. (2014). Not a victimless crime: The impact of fraud on individual victims and their families. *Security Journal*, 27(1), 36-54.
- Centraal Bureau voor de Statistiek. (2022, March 1). *Nearly 2.5 million people victims of cybercrime in 2021*. CBS. Retrieved November 14, 2022, from <https://www.cbs.nl/en-gb/news/2022/09/nearly-2-5-million-people-victims-of-cybercrime-in-2021>.
- Cialdini, R. B., & Cialdini, R. B. (2007). *Influence: The psychology of persuasion* (Vol. 55, p. 339). New York: Collins.
- Conway, A. R., & Engle, R. W. (1994). Working memory and retrieval: a resource-dependent inhibition model. *Journal of Experimental Psychology: General*, 123(4), 354.
- Cokely, E. T., & Kelley, C. M. (2009). Cognitive abilities and superior decision making under risk: A protocol analysis and process model evaluation.
- Coluccia, A., Pozza, A., Ferretti, F., Carabellese, F., Masti, A., & Gualtieri, G. (2020). Online romance scams: relational dynamics and psychological characteristics of the victims and scammers. A scoping review. *Clinical Practice and Epidemiology in Mental Health: CP & EMH*, 16, 24.
- Cui, X., Ge, Y., Qu, W., & Zhang, K. (2020). Effects of Recipient Information and Urgency

- Cues on Phishing Detection. In *HCI International 2020-Posters: 22nd International Conference, HCII 2020, Copenhagen, Denmark, July 19–24, 2020, Proceedings, Part III 22* (pp. 520-525). Springer International Publishing.
- Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of verbal learning and verbal behavior*, 19(4), 450-466.
- Darwish, A., El Zarka, A., & Aloul, F. (2012). Towards understanding phishing victims' profile. In *2012 International Conference on Computer Systems and Industrial Informatics* (pp. 1-5). IEEE.
- Dijker, A. J., & Koomen, W. (1996). Stereotyping and attitudinal effects under time pressure. *European Journal of Social Psychology*, 26(1), 61-74.
- Dodge, R., Coronges, K., & Rovira, E. (2012). Empirical benefits of training to phishing susceptibility. In *Information Security and Privacy Research: 27th IFIP TC 11 Information Security and Privacy Conference, SEC 2012, Heraklion, Crete, Greece, June 4-6, 2012. Proceedings 27* (pp. 457-464). Springer Berlin Heidelberg.
- Dong, X., Clark, J. A., & Jacob, J. L. (2008). User behaviour based phishing websites detection. In *2008 International Multiconference on Computer Science and Information Technology* (pp. 783-790). IEEE.
- Downs, J. S., Holbrook, M., & Cranor, L. F. (2007). Behavioral response to phishing risk. In *Proceedings of the anti-phishing working groups 2nd annual eCrime researchers summit* (pp. 37-44).
- Downs, J. S., Holbrook, M. B., & Cranor, L. F. (2006). Decision strategies and susceptibility to phishing. In *Proceedings of the second symposium on Usable privacy and security* (pp. 79-90).
- Duggan, G. B., & Payne, S. J. (2011). Skim reading by satisficing: evidence from eye tracking. In *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 1141-1150).
- Duggan, G. B., & Payne, S. J. (2009). Text skimming: The process and effectiveness of foraging through text under time pressure. *Journal of experimental psychology: Applied*, 15(3), 228.
- Eagly, A. H., & Chaiken, S. (1993). *The psychology of attitudes*. Harcourt brace Jovanovich college publishers.
- 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).

- Fette, I., Sadeh, N., & Tomasic, A. (2007). Learning to detect phishing emails. In Proceedings of the 16th international conference on World Wide Web (pp. 649-656).
- Finucane, M. L., Alhakami, A., Slovic, P., & Johnson, S. M. (2000). The affect heuristic in judgments of risks and benefits. *Journal of behavioral decision making*, 13(1), 1-17.
- Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic perspectives*, 19(4), 25-42.
- Ganzini, L., McFarland, B. H., & Cutler, D. (1990). Prevalence of mental disorders after catastrophic financial loss. *The Journal of nervous and mental disease*, 178(11), 680-685.
- Garera, S., Provos, N., Chew, M., & Rubin, A. D. (2007). A framework for detection and measurement of phishing attacks. In Proceedings of the 2007 ACM workshop on Recurring malcode (pp. 1-8).
- Hadnagy, C. (2018). *Social Engineering: The Science of Human Hacking*. Indianapolis, IN: Wiley.
- Halevi, T., Memon, N., & Nov, O. (2015). Spear-phishing in the wild: A real-world study of personality, phishing self-efficacy and vulnerability to spear-phishing attacks. *Phishing Self-Efficacy and Vulnerability to Spear-Phishing Attacks*.
- Harrison, B., Vishwanath, A., & Rao, R. (2016). A user-centered approach to phishing susceptibility: The role of a suspicious personality in protecting against phishing. In 2016 49th Hawaii international conference on system sciences (HICSS) (pp. 5628-5634). IEEE.
- Huang, L., Jia, S., Balcetis, E., & Zhu, Q. (2022). Advert: An adaptive and data-driven attention enhancement mechanism for phishing prevention. *IEEE Transactions on Information Forensics and Security*, 17, 2585-2597.
- Jagatic, T. N., Johnson, N. A., Jakobsson, M., & Menczer, F. (2007). Social phishing. *Communications of the ACM*, 50(10), 94–100.
- 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.
- Jones, H. S., Towse, J. N., & Race, N. (2015). Susceptibility to Email Fraud: A Review of Psychological Perspectives, Data-Collection Methods, and Ethical Considerations. *International Journal of Cyber Behavior, Psychology and Learning*, 5(3), 13–29.
- Jones, H. S. (2016). What makes people click: assessing individual differences in susceptibility to email fraud. Lancaster University (United Kingdom)

- Junglas, I., & Spitzmuller, C. (2006). Personality traits and privacy perceptions: an empirical study in the context of location-based services. In 2006 International Conference on Mobile Business (pp. 36-36). IEEE.
- Klein, C., & Ettinger, U. (Eds.). (2019). Eye movement research: An introduction to its scientific foundations and applications. Springer Nature.
- Kumaraguru, P., Rhee, Y., Acquisti, A., Cranor, L. F., Hong, J., & Nunge, E. (2007). Protecting people from phishing: the design and evaluation of an embedded training email system. In Proceedings of the SIGCHI conference on Human factors in computing systems (pp. 905-914).
- 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.
- Langenderfer, J., & Shimp, T. A. (2001). Consumer vulnerability to scams, swindles, and fraud: A new theory of visceral influences on persuasion. *Psychology & Marketing*, 18(7), 763-783.
- Linck, J. A., Osthus, P., Koeth, J. T., & Bunting, M. F. (2014). Working memory and second language comprehension and production: A meta-analysis. *Psychonomic bulletin & review*, 21, 861-883.
- Liu, Z., Zhou, L., & Zhang, D. (2020). Effects of Demographic Factors on Phishing Victimization in the Workplace. In PACIS (p. 75).
- Marett, K., & Wright, R. (2009). The effectiveness of deceptive tactics in phishing.
- Mayer, R., & Mayer, R. E. (Eds.). (2005). *The Cambridge handbook of multimedia learning*. Cambridge university press.
- Mayhorn, C. B., & Nyeste, P. G. (2012). Training users to counteract phishing. *Work*, 41(Supplement 1), 3549-3552.
- McAlaney, J., & Hills, P. J. (2020). Understanding phishing email processing and perceived trustworthiness through eye tracking. *Frontiers in Psychology*, 11, 1756.
- Mitnick, K. D., and Simon, W. L. (2002). *The Art of Deception*. Indianapolis: Wiley Publishing, Inc.
- Miyamoto, D., Iimura, T., Blanc, G., Tazaki, H., & Kadobayashi, Y. (2014). EyeBit: eye-tracking approach for enforcing phishing prevention habits. In 2014 Third International Workshop on Building Analysis Datasets and Gathering Experience Returns for Security (BADGERS) (pp. 56-65). IEEE.
- Modic, D., & Lea, S. E. (2012). How neurotic are scam victims, really? The big five and Internet scams. *The Big Five and Internet Scams*.

- Modic, D., & Lea, S. E. (2013). Scam compliance and the psychology of persuasion. Available at SSRN 2364464.
- Mohebzada, J. G., El Zarka, A., BHoiani, A. H., & Darwish, A. (2012). Phishing in a university community: Two large scale phishing experiments. In 2012 international conference on innovations in information technology (IIT) (pp. 249-254). IEEE.
- Myers, S. (2007). Introduction to phishing. In M. Jakobsson & S. Myers (Eds.), *Phishing and Countermeasures* (pp. 1–29). New Jersey: John Wiley & Sons, Inc.
- Nairne, J. S. (2010). Adaptive memory: Evolutionary constraints on remembering. In *Psychology of learning and motivation* (Vol. 53, pp. 1-32). Academic Press.
- Nicholson, J., Coventry, L. M., & Briggs, P. (2017). Can we fight social engineering attacks by social means? Assessing social salience as a means to improve phish detection. In *SOUPS* (pp. 285-298).
- Parsons, K., McCormac, A., Pattinson, M., Butavicius, M., & Jerram, C. (2015). The design of phishing studies: Challenges for researchers. *Computers & Security*, 52, 194-206.
- Paulhus, D. L., & Vazire, S. (2007). The self-report method. *Handbook of research methods in personality psychology*, 1(2007), 224-239.
- Petty, R. E., & Cacioppo, J. T. (1986). The elaboration likelihood model of persuasion. In *Communication and persuasion* (pp. 1-24). Springer, New York, NY.
- Pfeffel, K., Ulsamer, P., & Müller, N. H. (2019). Where the user does look when reading phishing mails—an eye-tracking study. In *Learning and Collaboration Technologies. Designing Learning Experiences: 6th International Conference, LCT 2019, Held as Part of the 21st HCI International Conference, HCII 2019, Orlando, FL, USA, July 26–31, 2019, Proceedings, Part I 21* (pp. 277-287). Springer International Publishing.
- Poole, A., & Ball, L. J. (2006). Eye tracking in HCI and usability research. In *Encyclopedia of human computer interaction* (pp. 211-219). IGI global.
- Ramkumar, N., Kothari, V., Mills, C., Koppel, R., Blythe, J., Smith, S., & Kun, A. L. (2020). Eyes on URLs: Relating visual behavior to safety decisions. In *ACM Symposium on Eye Tracking Research and Applications* (pp. 1-10).
- Redick, T. S., & Engle, R. W. (2006). Working memory capacity and attention network test performance. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, 20(5), 713-721.
- Redick, T. S., Heitz, R. P., & Engle, R. W. (2007). Working memory capacity and inhibition: Cognitive and social consequences.

- Rosenman, R., Tennekoon, V., & Hill, L. G. (2011). Measuring bias in self-reported data. *International journal of behavioural & healthcare research*, 2(4), 320.
- Shah, D. V., Kwak, N., Schmierbach, M., & Zubric, J. (2004). The interplay of news frames on cognitive complexity. *Human Communication Research*, 30(1), 102-120.
- Shah, J., Shah, R., & Liang, P. (2021). Too Much Nudging: Can it Cause a Decrease in the Desired Response?. *Journal of Student Research*, 10(4).
- Sheng, S., Holbrook, M., Kumaraguru, P., Cranor, L. F., & Downs, J. (2010). Who falls for phishing?: A demographic analysis of phishing susceptibility and effectiveness of interventions. *Proceedings of the 28th International Conference on Human Factors in Computing Systems - CHI '10*, 373.
- Sheng, S., Magnien, B., Kumaraguru, P., Acquisti, A., Cranor, L. F., Hong, J., & Nunge, E. (2007). Anti-phishing phil: the design and evaluation of a game that teaches people not to fall for phishing. In *Proceedings of the 3rd symposium on Usable privacy and security* (pp. 88-99).
- Stoet, G. (2010). PsyToolkit: A software package for programming psychological experiments using Linux. *Behavior research methods*, 42, 1096-1104.
- Stoet, G. (2017). PsyToolkit: A novel web-based method for running online questionnaires and reaction-time experiments. *Teaching of Psychology*, 44(1), 24-31.
- Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of experimental psychology*, 18(6), 643.
- Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases: Biases in judgments reveal some heuristics of thinking under uncertainty. *science*, 185(4157), 1124-1131.
- Valtakari, N. V., Hooge, I. T., Viktorsson, C., Nyström, P., Falck-Ytter, T., & Hessels, R. S. (2021). Eye tracking in human interaction: Possibilities and limitations. *Behavior Research Methods*, 1-17.
- Vayansky, I., & Kumar, S. (2018). Phishing—challenges and solutions. *Computer Fraud & Security*, 2018(1), 15-20.
- 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.
- Vitek, V., & Syed Shah, T. (2019). Implementing a Nudge to Prevent Email Phishing.

- Warkentin, M., McBride, M., & Carter, L. (2011). Dispositional and situational factors: influences on information security policy violations. *European Journal of Information Systems*, 25(3), 231-251.
- Weider, D. Yu, Shruti Nargundkar, and Nagapriya Tiruthani. (2008). "A phishing vulnerability analysis of web based systems." 2008 IEEE Symposium on Computers and Communications. IEEE, 2008.
- Workman, M. (2008). Wisecrackers: A theory-grounded investigation of phishing and pretext social engineering threats to information security. *Journal of the American Society for Information Science and Technology*, 59(4), 662e674.
- Wu, M., Miller, R. C., & Garfinkel, S. L. (2006). Do security toolbars actually prevent phishing attacks?. In *Proceedings of the SIGCHI conference on Human Factors in computing systems* (pp. 601-610).
- Xiong, A., Proctor, R. W., Yang, W., & Li, N. (2017). Is domain highlighting actually helpful in identifying phishing web pages?. *Human factors*, 59(4), 640-660.
- Yan, Z., & Gozu, H. Y. (2012). Online Decision-Making in Receiving Spam Emails Among College Students: *International Journal of Cyber Behavior, Psychology and Learning*, 2(1), 1-12.

Appendix A

Dear participant,

Thank you for considering to participate in this study!

It should take between 30-45 minutes to complete all the elements of this experiment. First, you will be asked to complete a short survey to establish demographic information, as well as information regarding your internet and email usage behaviour. Secondly, you will be requested to perform an email legitimacy task. With this task, your goal is to analyze a set of emails and assess their authenticity (real vs. fake). Your viewing behaviour will be monitored through a screen-based eye-tracker. More in-depth instructions of the task will be provided to you once administered. Lastly, after completing the email legitimacy task, you will be asked to perform three additional cognitive tasks. These tasks measure working memory capacity, inhibition and cognitive reflection. Detailed instructions of the cognitive tasks will be provided to you at the moment they will be administered.

Important to note is that your participation is completely voluntary and you are therefore able to withdraw at any time without providing a reason. Once you have decided to withdraw, all data that has been collected to that point will be deleted.

At any point during this study, your anonymity will be ensured. The received data is not personally identifiable and can therefore not be traced back to you. The unidentifiable data is used for academic purposes and might be shared with third parties in regards to sharing the results or publication of the study.

For additional information or questions about this study, please contact:

Jasper Rothert (Researcher): j.rothert@student.utwente.nl

Dr. Steven J. Watson (Supervisor): s.j.watson@utwente.nl

- I agree that I have fully read the information above.
- I agree to participate in this study and that my (unidentifiable) data will be used for academic purposes only.

Appendix B

Thank you for participating in this study, it is much appreciated!

In this survey, you will be asked to provide information about your internet and email usage habits. Please, try to fill in the questions as accurately as possible. First, some questions regarding your demographics will be asked, after which the internet and email usage habits questions will be shown.

You are requested to provide an answer to each question. The survey should not take longer than 5 minutes to complete.

Press the arrow in the bottom-right corner to continue to the questions.

Age

Gender

- Male
- Female
- Non-binary / third gender
- Prefer not to say

Nationality

- Dutch
- German
- Other, please indicate below

Highest level of finished education

- Elementary school
- High school
- Vocational education (MBO)
- Bachelor degree
- Master degree

For each of the questions below please choose the answer which best describes your internet and email usage habits.

How many hours do you spend actively using the internet on a typical day?

- 0-1 Hours
 - 1-3 Hours
 - 3-6 Hours
 - 6+ Hours
 - I do not use the internet on a daily basis
-

From your estimation, what proportion of this time is spent reading and responding to email correspondences?

- 0-20%
 - 20-40%
 - 40-60%
 - 60-80%
 - 80-100%
-

On a typical day, how many emails do you receive?

- 0-5 emails
 - 6-10 emails
 - 11-15 emails
 - 16-20 emails
 - 20+ emails
-

What kind of device do you use most frequently to read and respond to email correspondences? (Only pick a single device)

- Laptop/Desktop PC
 - Smartphone
 - Tablet
 - Other, please indicate below
-

To your knowledge, how many phishing emails do you receive in a typical week?

(Note: Phishing refers to a fraudulent communication sent from somebody that

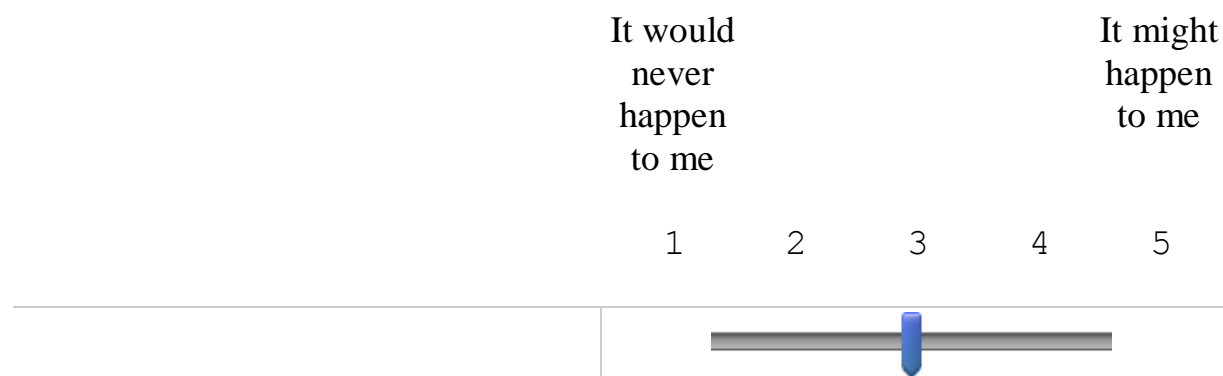
is posing to be someone else in order to elicit personal information from the user which can include for example usernames, passwords and bank account details, or to download and install an attachment that contains malware which can then be used to steal personal data from the user's computer.)

- 0-5 emails
 - 6-10 emails
 - 11-15 emails
 - 16-20 emails
 - 20+ emails
-

To your knowledge, have you ever responded to a phishing email?

- Yes
- No

Please indicate on the scale below how at risk you feel to online fraud. (You have to move the slider first before being able to proceed.)




Please indicate on the scale below your perceived knowledge to detect phishing emails. (You have to move the slider first before being able to proceed.)

I am not knowledgeable at all

I am highly knowledgeable

1 2 3 4 5



The image shows a horizontal slider scale with five numbered positions (1 to 5). The scale is bounded by two horizontal lines. A vertical line is positioned at the 1 mark. A grey horizontal bar represents the slider, starting from the 2 mark and extending to the 4 mark. A blue vertical slider knob is positioned at the 3 mark.

Appendix C

In this section of the study, you are requested to indicate on a scale of -3 to 3 whether you think each of the emails shown to you is either phishing or legitimate (-3 = definitely phishing, 3 = definitely legitimate, the values in between indicate how confident you are in your answer).

Phishing refers to a fraudulent communication sent from somebody that is posing to be someone else in order to elicit personal information from the user which can include for example usernames, passwords and bank account details, or to download and install an attachment that contains malware which can then be used to steal personal data from the user's computer.

For each of the emails that is shown to you during this task, you are to assume that all communications are relevant to the recipient, unless it is obvious otherwise. For instance, if an email in the task is from a social media platform, you are to assume that the recipient has an account at the platform the email is from. Please verbally indicate to the researcher on a scale from -3 to 3 after seeing each email how confident you are that the email on the screen is either phishing or legitimate. The scale will be shown to you on screen after each email.

First, you will perform a practice round of 2 emails to familiarize yourself with the task. With the first email, you have 8 seconds to view it before you have to give an answer. For the second email, you have 20 seconds. After the practice round, you will perform the actual email legitimacy task, in which you have to analyze 2 sets containing 16 emails each (32 emails total). For one set, you have 8 seconds viewing time per email and for the other set you will have 20 seconds. If you have any questions about the task, you can ask the researcher now. If not, you can press the space bar to continue to the practice round.

Appendix E

Practice trials (2 sentences):

In a flash of fatigue and fantasy, he saw a fat Indian sitting beside a campfire.
The lieutenant sat beside the man with the walkie-talkie and stared at the muddy ground.
I will not shock my readers with a description of the cold-blooded butchery that followed.
The courses are designed as much for professional engineers as for amateur enthusiasts.
It was shortly after this that an unusual pressure of business called me out of town.
He pursued this theme, still pretending to seek for information to quiet his own doubts.

2 sentences trials:

I was so surprised at this unaccountable apparition, that I was speechless for a while.
When at last his eyes opened, there was no gleam of triumph, no shade of anger.
Filled with these dreary forebodings, I fearfully opened the heavy wooden door.
I'm not certain what went wrong but I think it was my cruel and bad temper.
I imagine that you have a shrewd suspicion of the object of my early visit.
I turned my memories over at random like pictures in a photograph album.

3 sentences trials:

Sometimes I get so tired of trying to convince him that I love him and shall forever.
The woman hesitated for a moment to taste the onions because her husband hated the smell.
It was your belief in the significance of my suffering that kept me going.
When in trouble, children naturally hope for a miraculous intervention by a superhuman.
With shocked amazement and appalled fascination Marion looked at the pictures.
There are days when the city where I live wakes in the morning with a strange look.
We boys wanted to warn them, but we backed down when it came to the pinch.
He stood there at the edge of the crowd while they were singing, and he looked bitter.
What would come after this day would be inconceivably different, would be real life.

4 sentences trials:

John became annoyed with Karen's bad habits of biting her nails and chewing gum.
Due to his gross inadequacies, his position as director was terminated abruptly.
It is possible, of course, that life did not arise on the earth at all.
The poor lady was thoroughly persuaded that she was not long to survive this vision.

After all he had not gone far, and some of his walking had been circular.
 The announcement of it would resound throughout the world, penetrate to the remotest land.
 To do so in directions that are adaptive for mankind would be a realistic objective.
 Slicing it out carefully with his knife, he folded it without creasing the face.
 He laughed sarcastically and looked as if he could have poisoned me for my errors.
 He tolerated another intrusion and thought himself a paragon of patience for doing so.
 The reader may suppose that I had other motives, besides the desire to escape the law.
 On the desk where she wrote her letters was a clutter of objects coated in dust.

5 sentences trials:

He stuffed his denim jacket into his pants and fastened the stiff, new snaps securely.
 He had an odd elongated skull which sat on his shoulders like a pear on a dish.
 His imagination had so abstracted him that his name was called twice before he answered.
 The basic characteristic of the heroes in the preceding stories is their sensitivity.
 He listened carefully because he had the weird impression that he knew the voices.
 He had patronized her when she was a schoolgirl and teased her when she was a student.
 He covered his heart with both hands to keep anyone from hearing the noise it made.
 The stories all deal with a middle-aged protagonist who attempts to withdraw from society.
 Without tension there could be no balance either in nature or in mechanical design.
 I wish there existed someone to whom I could say that I felt very sorry.
 Here, as elsewhere, the empirical patterns are important and abundantly documented.
 The intervals of silence grew progressively longer; the delays became very maddening.
 Two or three substantial pieces of wood smoldered on the hearth, for the night was cold.
 I imagined that he had been thinking things over while the secretary was with us.
 There was still more than an hour before breakfast, and the house was silent and asleep.

6 sentences trials:

He sometimes considered suicide but the thought was too oppressive to remain in his mind.
 And now that a man had died some unimaginably different state of affairs must come to be.
 When I got to the big tobacco field, I saw that it had not suffered much.
 The products of digital electronics will play an important role in your future.
 One problem with this explanation is that there appears to be no defence against cheating.
 Sometimes the scapegoat is an outsider who has been taken into the community.
 I should not be able to make anyone understand how exciting it all was.

A small oil lamp burned on the floor and two men crouched against a wall, watching me.

The sound of an approaching train woke him, and he started to his feet.

The entire construction crew decided to lengthen their work day in order to have lunch.

The smokers were asked to refrain from their habit until the end of the production.

All students that passed the test were exempt from any further seminars that semester.

Despite the unusually cold weather, the campers continued their canoe trip.

The young business executive was determined to develop his housing projects within the year.

In order to postpone the business trip, he canceled his engagements for the week.

The incorrigible child was punished brutally for his lack of respect for his elders.

The brilliant trial attorney dazzled the jury with her astute knowledge of the case.

I found the keynote speaker incredibly boring, inarticulate and not well read.

Appendix F

Dear participant,

In this section of the research, you will be asked to answer three questions. Read the questions carefully. A response on each question is required and you should provide your answers in digits only (i.e., "2" instead of "Two"). After filling in all questions you can report back to the researcher. It should usually not take longer than 3 minutes to complete the entire task.

A bat and a ball cost €1.10 in total. The bat costs €1.00 more than the ball. How much does the ball cost? (Provide your answer in cents.)

If it takes 5 machines 5 minutes to make 5 widgets, how long would it take 100 machines to make 100 widgets? (Provide your answer in minutes.)

In a lake, there is a patch of lily pads. Every day, the patch doubles in size. If it takes 48 days for the patch to cover the entire lake, how long would it take for the patch to cover half the lake? (Provide your answer in number of days.)

Did you already know the answer to any of these questions before taking the test?

No

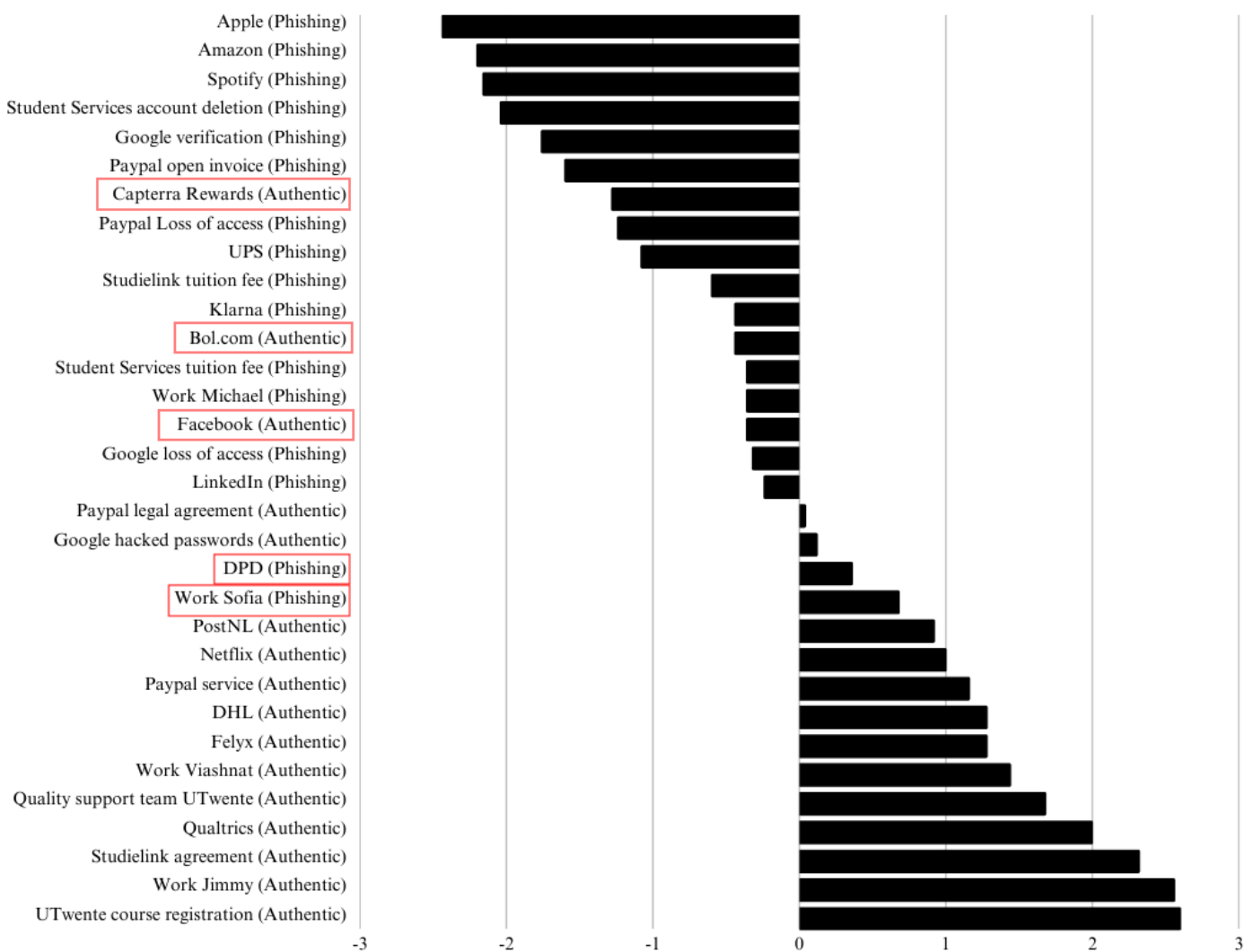
Yes

Appendix G

Figure 1 shows the average scores on the 6-point Likert scale for each email stimulus. Five emails (3 legitimate, 2 fraudulent) had mean responses located at the wrong side of the midpoint (indicated with red boxes), which indicates that they were more challenging to correctly classify, while 8 other stimuli (4 legitimate, 4 fraudulent) were easier to classify due to their means being near the extremes of the scale ($\geq |2|$) with a small standard deviation ($SD \leq 1.5$). This suggests that there was a diverse range in difficulty of the email stimuli used in the task.

Figure G1.

Diagram with mean scale rating across participants for each email stimulus. The red boxes around certain stimuli indicate that these were judged on the wrong side of the midpoint



Appendix H

Table H1:

Percentage responses to the questions of the email familiarity questionnaire

Question	0-1 hours	1-3 hours	3-6 hours	6+ hours	I do not use the internet on a daily basis.
1. How many hours do you spend actively using the internet on a typical day?	0	8.0	72.0	20.0	0
	0-20%	20-40%	40-60%	60-80%	80-100%
2. What proportion of this time is spent on reading and responding to email correspondences?	76.0	24.0	0	0	0
	Laptop/Desktop	Smartphone	Tablet	Other	
3. What kind of device do you use most frequently to read and respond to email correspondences?	60.0	40.0	0	0	
	0-5	6-10	11-15	16-20	20+
4. How many emails do you receive on a typical day?	44.0	40.0	8.0	8.0	0
5. How many phishing emails do you receive in a typical week	76.0	12.0	8.0	0	4.0
	Yes	No			

6. To your knowledge, have you ever responded to a phishing email?	24.0	76.0			
	1 (low)	2	3	4	5 (high)
7. Your perceived knowledge to detect phishing emails	0	24.0	20.0	52.0	4.0
8. How at risk do you feel to online fraud	12.0	56.0	12.0	16.0	4.0

Table H2.

Regression model including all questions of the email familiarity questionnaire as independent variables and the total number of correct answers on the email legitimacy task as the dependent variable.

Model	<i>B</i>	<i>SE</i>	β	<i>t</i>	<i>p</i>	Adj. <i>R</i> ²
(Constant)	24.85	7.60		3.27	.005	-.02
1. Daily time using internet	-0.49	1.45	-0.09	-0.33	.742	
2. Time reading/responding to emails	-0.40	1.45	-0.06	-0.27	.789	
3. Preferred device	-0.07	1.42	-0.01	-0.05	.962	
4. Number of emails per day	0.38	0.90	0.12	0.42	.678	
5. Number of phishing emails per week	-0.25	0.87	-0.09	-0.29	.774	
6. Responded to phishing emails	-2.72	1.61	-0.42	-1.69	.111	
7. Perceived knowledge phishing	1.20	0.83	0.39	1.45	.166	
8. Perceived risk online fraud	0.18	0.65	0.07	0.28	.785	

Table H3.

Regression model including perceived knowledge to detect phishing emails and perceived risk to online fraud as the independent variables and total number of correct answers on the email legitimacy task as the dependent variable

Model	<i>B</i>	<i>SE</i>	β	<i>t</i>	<i>p</i>	Adj. <i>R</i> ²
(Constant)	16.99	3.02		5.62	<.001	.08
7. Perceived knowledge phishing	1.30	0.65	0.42	2.00	.059	
8. Perceived risk online fraud	0.57	0.57	0.21	1.01	.323	

Appendix I

Table I1.

Fixation duration on the four phishing indicators by time condition.

		<i>M</i>	<i>SD</i>	<i>t</i>	<i>df</i>	<i>p</i>
Mismatched/Generic sender domain name	Time pressure	1.20	0.79	-8.27	23	<.001**
	No time pressure	3.19	1.71			
Suspicious links	Time pressure	0.38	0.36	-3.45	23	.002**
	No time pressure	0.90	0.70			
Urgency	Time pressure	0.63	0.32	-6.01	23	<.001**
	No time pressure	1.23	0.50			
Threat	Time pressure	0.57	0.36	-4.88	23	<.001**
	No time pressure	1.04	0.39			

***. Significant at the $p < .01$ level (2-tailed).*

**. Significant at the $p < .05$ level (2-tailed).*