Journal of Cybersecurity Education, Research and Practice

Volume 2021 | Number 2 Article 6

February 2022

EXPERIMENTAL STUDY TO ASSESS THE IMPACT OF TIMERS ON USER SUSCEPTIBILITY TO PHISHING ATTACKS

Amy E. Antonucci
Western Governors University, aa2539@mynsu.nova.edu

Yair Levy
Nova Southeastern University, USA, levyy@nova.edu

Laurie P. Dringus
Nova Southeastern University -- College of Engineering and Computing, laurie@nova.edu

Martha Snyder
smithmt@nova.edu

EXPERIMENTAL STUDY TO ASSESS THE IMPACT OF TIMERS ON USER SUSCEPTIBILITY TO PHISHING ATTACKS

Abstract

Social engineering costs organizations billions of dollars. It exploits the weakest link of information systems security, the users. It is well-documented in literature that users continue to click on phishing emails costing them and their employers significant monetary resources and data loss. Training does not appear to mitigate the effects of phishing much; other solutions are warranted. Kahneman introduced the concepts of System-One and System-Two thinking. System-One is a quick, instinctual decision-making process, while System-Two is a process by which humans use a slow, logical, and is easily disrupted. The key aim of our experimental field study was to investigate if requiring the user to pause by presenting a countdown or count-up timer when a possible phishing email is opened will influence the user to enter System-Two thinking. In this study, we designed, developed, and empirically tested a Pause-and-Think (PAT) mobile app that presented a user with a warning dialog and a countdown or count-up timer. Our goal was to determine whether requiring users to wait with a colored warning and a timer has any effect on phishing attempts. The study was completed in three phases with 42 subject matter experts and 107 participants. The results indicated that a countdown timer set at 3-seconds accompanied by red warning text was most effective on the user's ability to avoid clicking on a malicious link or attachment. Recommendations for future research include enhancements to the PAT mobile app and investigating what effect the time of day has on susceptibility to phishing.

Keywords

Cybersecurity mitigation, phishing in mobile devices, timers in cybersecurity, social engineering, judgment error in cybersecurity, phishing email mitigation

INTRODUCTION

Social engineering has demonstrated to be one of the top threats to organizations causing financial damages the have reached billions of dollars every year (FBI, 2018). Social engineering is a technique in which the attacker attempts to build a relationship with the victim to convince the victim to give the attacker information or to perform other actions that lead to malicious impact or financial losses (Krombholz et al., 2015). Phishing is an example of a social engineering attack. Phishing is an e-mail- or instant-messaging-based attack aimed at a large group in which the attacker attempts to convince the intended victim to take some action such as click on a malicious link. Attackers use phishing to create a fear response in their victims that leads them to use heuristics eventually leading to systematic errors or ransomware (Goel et al., 2017; Kahneman, 2011).

Kahneman (2011) referred to the process by which humans use heuristics to make a quick decision as System One. System One is a quick, instinctual decision-making process. Examples of System One processes are orienting to a sudden sound or an experienced driver pressing the brake when faced with road danger. In contrast, Kahneman (2011) identified the process by which humans use a slow, logical process as System Two. System Two requires attention, is much slower, and is easily disrupted. Examples of System Two are looking for a person with a certain characteristic or checking the validity of a complex logical argument. Text color can also affect user judgement (Kahneman, 2011). Anderson et al. (2015) stated that text color in a warning message should stand out to the user so that the user's attention is captured. Poor user judgement leads to extraordinary monetary and data costs.

In this study, we are attempting to address the massive costs that social engineering impact on organizations (FBI, 2018; Musuva et al., 2019; Salahdine & Kaabouch, 2019). Since social engineering is such a significant financial problem, it is of interest to study what can be done to mitigate it. This study focuses on the problem of why users make judgement errors when evaluating the risks involved in clicking on an unknown link in an e-mail. Even when warned, users choose to put aside security concerns when deciding whether or not to follow links presented in an e-mail (Vance et al., 2018). A possible explanation for this is that users do not properly evaluate the risk involved in clicking on an unknown link, especially when overworked (Bravo-Lillo et al., 2011). Hirshleifer et al. (2019) found that financial analysts produce better forecasts when they are not mentally fatigued and use heuristics as they get more fatigued. Tversky and Kahneman (1974) stated that heuristics are assumptions made to simplify decisions and that users can be taught to recognize when they are using heuristics to decide. By requiring the user to pause in this study, the user's thought stream may be interrupted, and the user may be switched to System Two thinking.

In addition to the fact that repetitive messaging appears to disengage users, the color of a message also appears to help or hinder user attention (Kahneman, 2011). Wogalter et al. (2002) stated that red has been found to increase the hazard rating of a warning, and that colored labels, especially red, are more noticeable than grey. Anderson et al. (2015) found no difference in user attention when a warning was presented in red rather than grayscale. They acknowledged that their finding was contrary to prior research and encouraged further research on the topic of warning text color. Using text color to digitally nudge the user may increase the likelihood of capturing the user's attention. Thus, the main goal of this research study was to determine through experimental field study whether requiring e-mail users to pause by displaying a colored warning (grey, red, or black text) with a timer (countdown or count-up) when they are presented with a potentially malicious link has any effect on the percentage of users falling to phishing attempts. Our study first validated the experimental procedures using Subject Matter Expert (SMEs) panel. Additionally, the study addressed the following five research questions:

RQ1: What are the three timer values to require the user to pause that should be used in this experimental field study to assess users' ability to identify malicious links in e-mail according to cybersecurity SMEs?

RQ2: What level of functional correctness and validity of the custom-designed mobile app is sufficient according to cybersecurity SMEs?

RQ3: Are there statistically significant mean differences between the ability to avoid clicking on a malicious link of e-mail users who are not required to pause vs. email users who are required to pause at three separate timer values displayed with a warning in (a) grey, (b) red, or (c) black warning text?

RQ4: Are there statistically significant mean differences between the ability to avoid clicking on a malicious link of e-mail users who are not required to pause vs. email users who are required to pause at three separate timer values displayed with: (a) countdown timer, (b) count-up timer, or (c) no timer?

RQ5a: Are there statistically significant mean differences between the ability to avoid clicking on a malicious link of e-mail users who are not required to pause vs. email users who are required to pause at three separate timer values displayed with a warning in (a) grey, (b) red, or (c) black warning text based on the categories of: (a) age, (b) gender, (c) education level, (d) attention span, and (e) the volume of email that the user receives in a day?

RQ5b: Are there statistically significant mean differences between the ability to avoid clicking on a malicious link of e-mail users who are not

required to pause vs. email users who are required to pause at three separate timer values displayed with: (a) countdown timer, (b) count-up timer, or (c) no timer based on the categories of: (a) age, (b) gender, (c) education level, (d) attention span, and (e) the volume of email that the user receives in a day?

LITERATURE REVIEW

This section includes a literature review of social engineering, phishing, heuristics, security in mobile devices, phishing mitigation techniques, timers, and text color. There is some disagreement whether training or different text colors are useful to mitigate phishing. It's generally agreed that security in mobile devices is still an open problem and that the study of heuristics can help to mitigate phishing.

Social Engineering

Social engineering is one of the most under-researched, however, most effective cybercrimes (Jain et al., 2016). Social engineering is defined as "the art of exploiting the weakest link of information security systems: the people who are using them" (Jain et al., 2016, p. 94). Mihelič et al. (2019) called the human factor in social engineering a lever that is exploited by attackers. There are four stages of social engineering: (1) information gathering, (2) gain trust, or hook relationship, (3) exploit trust and execute attack, and (4) exit (Mitnick & Simon, 2003; Salahdine & Kaabouch, 2019). In the information gathering stage, the attacker performs a reconnaissance, which is an information gather about their target. In the hook relationship phase, the attacker baits the victim with fear or excitement (Goel et al., 2017). In the play exploitation and execution phase, the attacker executes the attack, and in the out phase, the attacker leaves with no or limited trace that they were ever there. Technical solutions to combat social engineering typically do not work (Krombholz et al., 2015), and Jain et al. (2016) said that there are no technical solutions to the problem of social engineering. Users are often too confident in their ability to detect a social engineering attack (Krombholz et al., 2015), partially because social engineers are becoming more devious. This means that methodologies for countering social engineering that were suggested just two years ago are no longer useful.

Phishing

While phishing is only one of 20 different kinds of social engineering defined by Salahdine and Kaabouch (2019), they stated that phishing is the most common type

of social engineering attack. Thompson (2012) stated that many attacks start with a bad user decision, i.e. human judgement error, and that anyone can be tricked by a phishing attack. A number of studies presented a variety of taxonomies (Gupta et al., 2018; Rastenis et al., 2020; Salahdine & Kaabouch, 2019). Salahdine and Kaabouch (2019) organized phishing attacks into five categories: spear, whaling, vishing, interactive voice response, and business email compromise while Rastenis et al. (2020) gave a wider definition, which included the devices and other media used. Gupta et al. (2018) offered a taxonomy based on the phases of a phishing attack. A number of studies focused on spear-phishing (Burns et al., 2019; Butavicius et al., 2015; Halevi et al., 2015; Hanus et al., 2021; Mihelič et al., 2019; Oliveira et al., 2017), and all of the studies ran a simulated phishing campaign. Hanus et al. (2021) used machine learning to predict who would be a victim to phishing. They found that spear phishing is more likely to successfully phish the user, and they found that many demographic factors have bearing on phishing victimization. They also found that the amount of attention that a user can devote to identifying a phish is significant.

Heuristics

In his book *Thinking Fast and Slow*, Kahneman (2011) introduced the concepts of System One and System Two as methods of describing human cognition. System One represents an instinctual thought process that comes quickly and automatically and requires little or no effort. Examples of System One are the ability to orient to a sudden sound or to detect if one object is closer than another (Kahneman, 2011). System Two is a slow, methodical thought process that requires deliberate effort. Examples of System Two are solving a complex mathematical equation or monitoring one's behavior in a social situation (Kahneman, 2011). A third model of decision making called the Recognition-Primed Decision (RPD) model was introduced by Klein (1993) and used by Rosa et al. (2021). Klein (1993) described the RPD model as a model in which the decision maker does not make a choice between two or more options, but instead acts based on prior experience. Klein (1993) used the example of a firefighter chief in action at a fire. Asked afterwards how he chose what to do, the chief stated that he made no conscience choice and simply sprang into action (Klein, 1993).

Security in Mobile Devices

When compared to phishing using a desktop computer, phishing using a mobile device has not been widely studied (Bottazzi et al., 2015; Mukhopadhyay & Argles, 2011). Challenges unique to a mobile device platform include a smaller screen which leads users not to see certain phishing cues that they might in a larger screen (Goel & Jain, 2018; Ndibwile et al., 2019) and which requires that some browser features be eliminated, including anti-phishing security features (Ndibwile et al.,

2017; Virvilis et al., 2014). Universal Resource Locators (URLs) are usually hidden by default in a mobile browser, decreasing user attention to any phishing cues in the URL (Chorghe & Shekokar, 2016). Users do not give as much attention to cues in mobile device browsers as they do in desktop browsers because of the smaller screen (Amro, 2018). Users also tend to trust their mobile device because their device is usually close to them (Amro, 2018). Goel and Jain (2018) discussed the security challenge of the physical mobile device, which typically has additional vulnerabilities such as a camera, the user's physical location, and access to SMS.

Phishing Mitigation Techniques

A polymorphic dialog is one that changes in appearance each time it displays and has been used in an attempt to mitigate phishing (Anderson et al., 2016; Brustoloni & Villamarín-Salomón, 2007). Overall, all the studies in this section found polymorphic warnings to be more effective than static warnings (Brustoloni & Villamarín-Salomón, 2007; Vance et al., 2018). Training is another common method of phishing mitigation. There is disagreement regarding whether anti-phishing training is effective. Burns et al. (2019), Goel and Jain (2018), as well as Junger et al. (2017) found anti-phishing training to be largely ineffective. Kumaraguru (2009), Sun et al. (2017), and Volkamer et al. (2018) found anti-phishing training to be largely effective. This disagreement suggests that anti-phishing training as it is implemented today may not be effective, but that a solution that uses components of anti-phishing training may be useful.

Timers

Few studies were found regarding social engineering that employed timers. Molinaro (2019) used a countdown timer during which her participants were asked to distinguish phishing e-mails from valid e-mails, but the timer was not the focus of her study. However, work related to timers in other research fields, namely healthcare, civil engineering, and psychology, have been conducted. In the field of healthcare, the research showed that timers are used to remind workers of a task or of a medical emergency. Marto et al. (2016) found that introducing a countdown timer with a reminder that stroke is an emergency to an emergency stroke patient's room decreased the time between when the patient arrived in the emergency room and the time the patient received a drug that is able to dissolve a clot. Lindahl et al. (2019) created an Android tablet app that allows patients to self-administer a blood-pressure test. In the app, the timer reminded the patient to sit still for five minutes. Lindahl et al. (2019) reported that 99% of 100 pregnant women followed the timer guidance and were able to complete the blood-pressure test. Hung et al. (2020) created a smartphone app to guide hospital cleaning staff in the cleaning of patient beds. The app alerted staff to which beds needed to be cleaned and provided a countdown timer to indicate the deadline for cleaning the bed. Hung et al. (2020)

stated that there was a significant decrease in time required for cleaning beds when the app was in use. The civil engineering literature regarding timers investigated Pedestrian Countdown Signals (PCS) at intersections. A PCS is a countdown timer that indicates to a pedestrian waiting to cross a road at an intersection when it is safe to cross (Keegan & O'Mahony, 2003). Biswas et al. (2017) studied the effect PCS and Driver Countdown Signals (DCS) had on the interaction between drivers and pedestrians. They found that the number of drivers that drove through a red light increased when a DCS was present, and that as the DCS neared zero, drivers moved into the crosswalks, blocking pedestrian movement. They concluded that PCS and DCS have an overall positive effect on traffic flow but an overall negative effect on pedestrian safety. Many areas of psychology have been represented by studies that include timers including somnology (Lo et al., 2019), urgent decision making (Cheong, 2018), standardized testing (Brooks et al., 2003), child psychology (Newquist et al., 2012), and remote team communication (Fine, 2016).

Text Color

There appears that very limited research exists that investigated the effect of text color in phishing warning notices. Anderson et al. (2015) investigated the effect of color warning images versus greyscale warning images, and other studies investigated text color, but not in the cybersecurity field (Silver et al., 2002; Wogalter et al., 2002). There are inconsistencies with regard to the effect of text color on the hazard perception of a warning. Wogalter et al. (2002) stated that red has been found to increase the hazard rating of a warning, and that colored labels, especially red, are more noticeable than grey.

METHODOLOGY

Overview of Research Design

This research was conducted in three phases as shown in Figure 1. It was hypothesized that the Pause and Think (PAT)^{TM} mobile app would help users to detect phishing by displaying a warning dialog in colored text and with a timer to move them into a more logical thought process. In Phase I, quantitative approaches were used to collect SME opinion on the value for the countdown or count-up timer in the warning dialog, on the validity of the sample e-mails, and on the experimental procedures of PAT. PAT was designed and developed during Phase II. Phase III used a quantitative approach to collect data from users using the app.

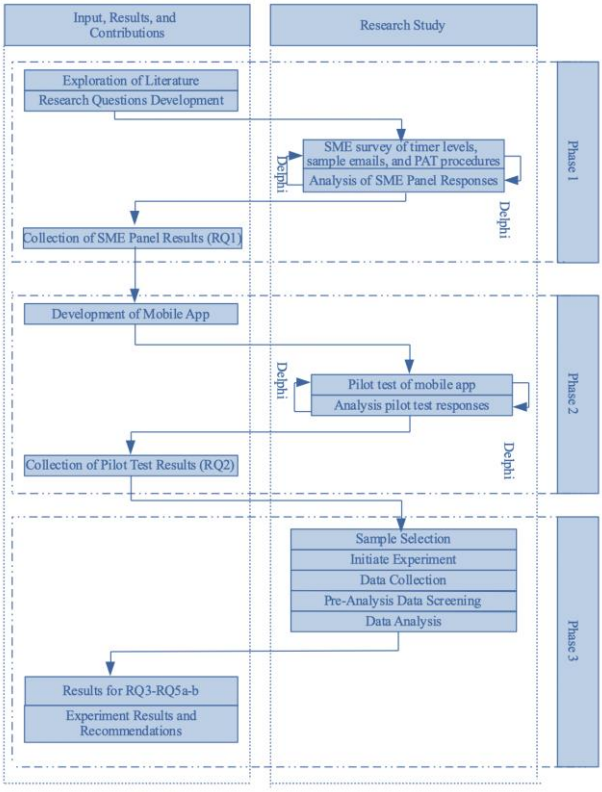

Figure 1. *Overview of Research Design Process for PAT*

Phase I – Expert Panel Validations

In Phase I, a four-section quantitative survey was used to collect opinion data from 42 SMEs on which timer value should be used in the countdown and count-up timer (Section 2), on the validity of the sample e-mails (Section 3), and on the experimental procedures of PAT (Section 4). The first section was a demographic questionnaire to document the expertise of the SME. The second section included a mockup of the timer dialog within the app so that the SMEs could visualize the process. This mockup is shown in Figure 2. The SMEs were then asked to rank eight timer values. After the data were collected, a second Delphi process round was completed to gain a more valid consensus (Ramim & Lichvar, 2014). The third section asked the SMEs to rate thirty sample emails individually. For each sample email, the SMEs were asked to identify the email as phishing or legitimate and whether the email should be kept, adjusted, or replaced. If the SME chose the option to adjust or replace, they were asked to specify how (in the case of adjust), or why (in the case of replace). The SMEs were also asked for additional feedback.

Phase II – PAT Mobile App Design and Development

Phase II entailed the design, development, and testing of PAT. PAT was created twice, once for Android devices and once for Apple devices. PAT simulates a basic Gmail client that allows the user to check their e-mail. PAT includes a demographic survey that is displayed the first time the app is opened. A warning and a timer as shown in Figure 2 displayed each time the user receiving the treatment opened a simulated e-mail that contained a URL or attachment. The user was not able to bypass the timer and had to wait until the timer was expired before interacting with the simulated e-mail. Each time the user interacted with a simulated e-mail for which a timer displayed, the id of that e-mail and whether the user clicked on the URL or attachment was stored.

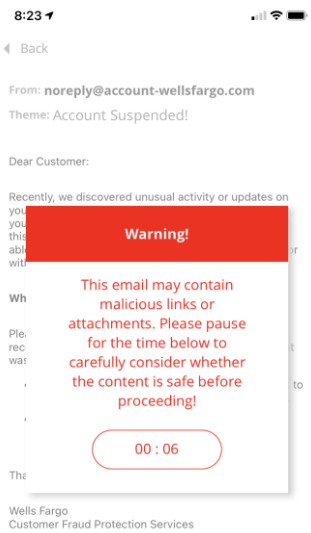

Figure 2. *Example of PAT Timer Dialog*

Phase III – Mobile App Delivery

In phase III, 107 participants, who were recruited via Facebook and LinkedIn, were asked to interact with PAT. The first 10 participants were recruited for a pilot study such that Apple and Android users were equally represented. The pilot group was used to verify the mobile app and data collection. The pilot participants were asked for feedback regarding the app and the findings and recommendations of the participants in the pilot study were incorporated into the app. Yan et al. (2015) studied user behavior for one week. Since this study was also analyzing user behavior, participants were asked to use PAT for seven days. For this study, simulated emails were randomly assigned to all participants from a pool of all

emails stored in the back-end. The pool contained ten legitimate text-only emails, and five each of the following: legitimate with a link, legitimate with an attachment, phishing with a link, and phishing with an attachment. Each participant received the same simulated emails each day, and each participant received five simulated emails per day. PAT collected and stored non-PII data from the participants. When participants downloaded PAT, they were given a User Identification Number (UIN) which was used to link their data to their profile. The participants were asked to take a short survey which included demographic questions. Participant age, gender, education level, attention span, and the amount of email they receive was stored. The survey also asked whether the participant is completely color-blind (National Eye Institute, 2019). Attention span was measured with an attention span test adopted from Psychology Today (n.d.) which was embedded in the app survey. After the participants finished the survey, the participants saw a simulated inbox listing. Participants were able to interact with any e-mail in the simulated inbox as though it were a real e-mail. The app had pre-coded simulated e-mails that displayed in a random order. Some of the simulated e-mails mimicked a legitimate e-mail, and some simulated a phish, and each simulated e-mail was identified by an id. New e-mails displayed on each day of the study to simulate receiving new e-mail. Some simulated e-mails had a URL or an attachment, and some did not. If a participant receiving the timer treatment opened an e-mail that had a URL in the body of the message or an attachment, a timer was displayed, and they were not able to interact with the e-mail until the timer expired. When they did interact with the email, the data collected was: (1) the ID of the simulated e-mail, and (2) whether the participant clicked on the link or attachment. The app also captured and stored whether a countdown, count-up, or no timer was used, the value of the timer used, and whether grey, red, or black text was used.

FINDINGS

This study resulted in the design and development of a mobile app called PAT which was used to test user reaction to a timer which was presented when the user opened an email that contains a link or attachment. The purpose of the timer was to assist the user's ability to avoid clicking on a malicious link or attachment. It was found that the timer did appear to assist the user's ability to avoid clicking on a malicious link or attachment.

Phase I – SME Survey Feedback and Findings

To answer RQ1 and RQ2, a group of 42 cybersecurity experts participated in Phase I of the study. One third of the SMEs (14) had at least three years of cybersecurity experience, and one third had two cybersecurity certifications. For the timer section of the survey, values of 3-seconds, 5-seconds, and 7-seconds were chosen, answering RQ1, which were then coded for use in the app in Phases II and III. In

the verification of sample emails section of the survey, of 10 sample phishing emails, most SMEs correctly identified only one phishing sample email as phishing. Many of the phishing sample emails were adjusted or replaced based on SME quantitative feedback. Of 20 legitimate sample emails, most SMEs correctly identified 14 legitimate sample emails as legitimate. Most SMEs recommended keeping all sample emails. In the mobile app experimental procedure section of the SME survey, SMEs were asked whether major components of the PAT process should be kept, adjusted, or removed. Most SMEs recommended keep for all of the components of PAT. These results answered RQ2.

Phase II – PAT Mobile App Development

Phase II consisted of the development of PAT. The development of PAT used SME feedback on timer value, sample email verification, and the mobile app experimental procedures. PAT was tested and deployed to both the Apple Store and Google Play. After the participants registered and logged in for the first time, they were asked demographic questions that included, age, gender, education level, volume of email, and a set of five questions designed to capture the value of the participant's attention span. The survey also asked if the participant was color-blind. Any participants that indicated that they were color-blind were excluded from the study. When the participant logged in at least one day after registering, a simulated inbox was displayed in the app. Simulated emails were coded based on SME feedback in Phase I. When a participant in the experimental group tapped on an email with a link or attachment, the simulated email opened and a timer dialog was displayed, as shown in Figure 2. After the timer dialog self-dismissed, if the participant tapped on the link, an acknowledgement of the tap was displayed.

Phase III – PAT Mobile App Delivery

Phase III involved participant download, installation, and use of PAT. A total of 117 participants downloaded the PAT mobile app and participated in the study. Of the 117 participants who participated, 10 were pilot testers. Five each of the pilot testers were Apple and Android users. Each tester was given a list of actions to take with the app. Each tester met with the researcher in person or online and the researcher watched the tester use the app. Minor issues were found and fixed. Other than the pilot testers, 107 users participated in the study. One user indicated that they were completely color blind. The results from that user were excluded from the study. The total remaining number of participants was 106. Any email interaction records that indicated that the participant did not open the email were excluded from the study. The number of email interactions collected was 3,746 (106 participants interacting with five emails per day for seven days on average). The data were filtered to include only email interactions with the simulated

phishing emails for a remaining total of 1,796 email interactions. The 106 participants included several demographic characteristics. Of the participant ages, one was 18-19 (0.93%), 13 were 26-35 (12.04%), 28 were 36-45 (25.93%), 40 were 46-55 (37.04%), 18 were 56-65 (16.67%), five were 66-75 (4.63%), and one was over 75 (0.93%). Of participant genders, 70 were female (64.81%) and 36 were male (33.33%). Of education level, no participants had a Below High School education and two (1.85%) had a High School education. Eleven (10.19%) participants had Some Higher Education Credits, six (5.56%) had an Associate Degree, 27 (25.00%) had a Bachelor Degree, 41 (37.96%) had a Master Degree, and 19 (17.59%) had a Doctorate Degree or comparable. Of volume of email, eight (7.41%) had 1-10 emails per day, 32 (29.63%) had 11-30 emails per day, 27 (25.00%) had 31-60 emails per day, 19 (17.59%) had 61-90 emails per day, eight (7.41%) had 91-120 emails per day, 5 (4.63%) had 121-150 emails per day, and seven (6.48%) had more than 150 emails per day. Attention span was aggregated from the five attention span demographic survey questions so that a lower score means a lower attention span. Each question was scored and added so that the minimum score was five, meaning that the participant scored the lowest attention span choice in each of the five questions. The maximum score was 33, which means that the highest-scoring participant scored two fewer than the possible maximum of 35 (five questions times a score of seven per question). The range of scores was then grouped so that scores of five through eight were coded as Very low attention span, scores of nine through 12 were coded as Low attention span, scores of 13 through 16 were scored as Somewhat low attention span, scores of 17 through 20 were scored as Average attention span, scores of 21 through 24 were scored as Somewhat high attention span, scores of 25 through 28 were scored as High attention span, and scores of 29 through 33 were scored as Very high attention span.

Phase III addressed RQ3. To answer RQ3, 3,746 email interactions were collected (106 participants interacting with 5 emails per day for 7 days on average). The data were filtered to include only email interactions with the simulated phishing emails. ANOVA was used to test for significant differences between groups. The results of the ANOVA showed there were significant differences among all groups for Text Color, Timer Value, and Text Color x Timer Value. The F-value for Text Color was 20.852 and had a significance of $p < 0.001$. The F-value for Timer Value was 3.700 and had a significance of $p < 0.05$. The F-value for Text Color x Timer Value was 2.899 and had a significance of $p < 0.01$. The results of the ANOVA to answer RQ3 are shown in Table 1.

Source	Sum of Squares	Df	Mean Square	F	Sig.
Text Color vs Timer Value					
Text Color	6.051	2	3.025	20.852	***0.000****
Timer Value	1.611	3	0.537	3.700	***0.011****
Text Color x Timer Value	2.524	6	0.421	2.899	***0.008****
Timer Type vs Timer Value					
Timer Type	0.049	1	0.049	0.328	0.567
Timer Value	0.655	2	0.327	2.207	0.110
Timer Type x Timer Value	1.039	2	0.520	3.501	***0.030****
Timer Type	0.049	1	0.049	0.328	0.567

* p <0.05, ** p <0.01, *** p <0.001

Table 1. *ANOVA Results (N=1796)*

The profile plot of Text Color x Timer Value is shown in Figure 3(a). The value of the Estimated Marginal Means of Clicked range from one, meaning Not Clicked, to two, meaning Clicked. The black line indicates the mean click rate for email interactions that included a dialog box in black text. Likewise, the grey line represents the mean click rate for email interactions that included a dialog box in grey text, and the red line indicates the mean click rate that included a dialog box in red text. The profile plot indicates that grey and red text performed better overall than black text, meaning that the user was less likely to click on a malicious link if the text color was in grey or red. The profile plot shows that the best combination of text color and timer value was grey text at 7-seconds. This combination had the lowest click mean at 1.65. The second-best combination was red text at 3-seconds. The click mean for this combination was approximately 1.67.

Phase III also addressed RQ4. To answer RQ4, the data were filtered to include only email interactions with the simulated phishing emails. ANOVA was used to test for significant differences between groups. The results of the ANOVA showed there were significant differences only in the Timer Type x Timer Value group. The F-value for Timer Type x Timer Value was $p < 0.05$. The results of the ANOVA to answer RQ4 are shown in Table 1.

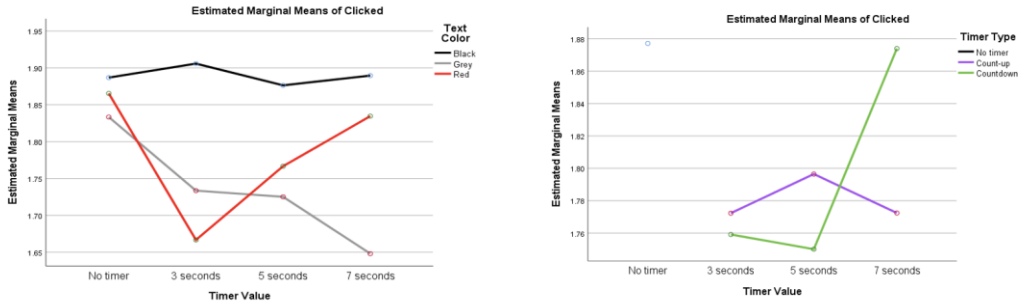

Figure 3. *Profile Plot of (a) Text Color x Timer Value and (b) Timer Type x Timer Value*

The profile plot for Timer Type x Timer Value is shown in Figure 3(b). No timer is represented by only a dot because there was no timer value for dialogs with no timer. The worst combinations of Timer Type and Timer Value were no timer and no time and a countdown timer at 7-seconds, both at a mean click rate of approximately 1.88. The best combination of Timer Type and Timer Value was a timer counting down for 5-seconds at a mean click rate of approximately 1.75.

Phase III addressed RQ5a. To answer RQ5a, the data were filtered to include only email interactions with the simulated phishing emails. ANCOVA was used to test for significant differences between groups with each demographic indicator as a covariate. The results of ANCOVA using all five demographic indicators (age, gender, education level, email volume, and attention span) showed significance. The results of the ANCOVA answering RQ5a are shown in Table 2.

Source	Sum of Squares	Df	Mean Square	F	Sig.
Age	0.135	1	0.135	0.931	0.335
Text Color	5.770	2	2.885	19.884	***0.000******
Timer Value	0.135	1	0.135	0.931	0.335
Text Color x Timer Value	2.428	6	0.405	2.789	***0.011****
Timer Type	0.035	1	0.035	0.234	0.629
Timer Value	0.642	2	0.321	2.167	0.115
Timer Type x Timer Value	1.04	2	0.52	3.51	***0.030****
Gender	0.027	1	0.027	0.185	0.667
Text Color	6.050	2	3.025	20.841	***0.000******
Timer Value	1.613	3	0.538	3.703	***0.011****

Text Color x Timer Value	2.545	6	0.424	2.923	***0.008***
Timer Type	0.049	1	0.049	0.333	0.564
Timer Value	0.654	2	0.327	2.204	0.111
Timer Type x Timer Value	1.04	2	0.52	3.502	***0.030****
Education Level	2.093	1	2.093	14.533	***0.000****
Text Color	6.101	2	3.051	21.185	***0.000****
Timer Value	1.810	3	0.603	4.191	***0.006***
Text Color x Timer Value	2.257	6	0.376	2.612	***0.016****
Timer Type	0.033	1	0.033	0.226	0.634
Timer Value	0.652	2	0.326	2.213	0.110
Timer Type x Timer Value	1.027	2	0.513	3.486	***0.031****
Email Volume	.960	1	0.960	6.641	***0.010****
Text Color	6.074	2	3.037	20.998	***0.000****
Timer Value	1.607	3	0.536	3.705	***0.011****
Text Color x Timer Value	2.547	6	0.424	2.935	***0.007***
Timer Type	0.048	1	0.048	0.327	0.567
Timer Value	0.679	2	0.339	2.294	0.101
Timer Type x Timer Value	1.041	2	0.520	3.517	***0.030****
Attention Span	0.023	1	0.023	0.160	0.690
Text Color	6.042	2	3.021	20.813	***0.000****
Timer Value	1.626	3	0.542	3.733	***0.011****
Text Color x Timer Value	2.523	6	0.421	2.897	***0.008***
Timer Type	0.049	1	0.049	0.327	0.567
Timer Value	0.654	2	0.327	2.204	0.111
Timer Type x Timer Value	1.036	2	0.518	3.491	***0.031****

* p <0.05, ** p <0.01, *** p <0.001

Table 2. *ANCOVA Results (N=1796)*

Profile plots of Text Color x Timer Value with each covariate were performed and appear in Figure 4. Figure 4(a) shows the profile plot of Text Color x Timer Value with age as a covariate. Figure 4(b) shows the profile plot of Text Color x Timer Value with gender as a covariate. Figure 4(c) shows the profile plot of Text Color x Timer Value with education level as a covariate. Figure 4(d) shows the profile plot of Text Color x Timer Value with email volume as a covariate. Figure 4(e) shows the profile plot of Text Color x Timer Value with attention span as a covariate. Phase III addressed RQ5b. To answer RQ5b, the data were filtered to include only email interactions with the simulated phishing emails. ANCOVA was used to test for significant differences between groups with each demographic

indicator as a covariate. The results of ANCOVA using all five demographic indicators (age, gender, education level, email volume, and attention span) showed significance. F-value for Timer Type x Timer Value was $p < 0.05$ for all demographic factors. The results of the ANCOVA answering RQ5a are shown in Table 2.

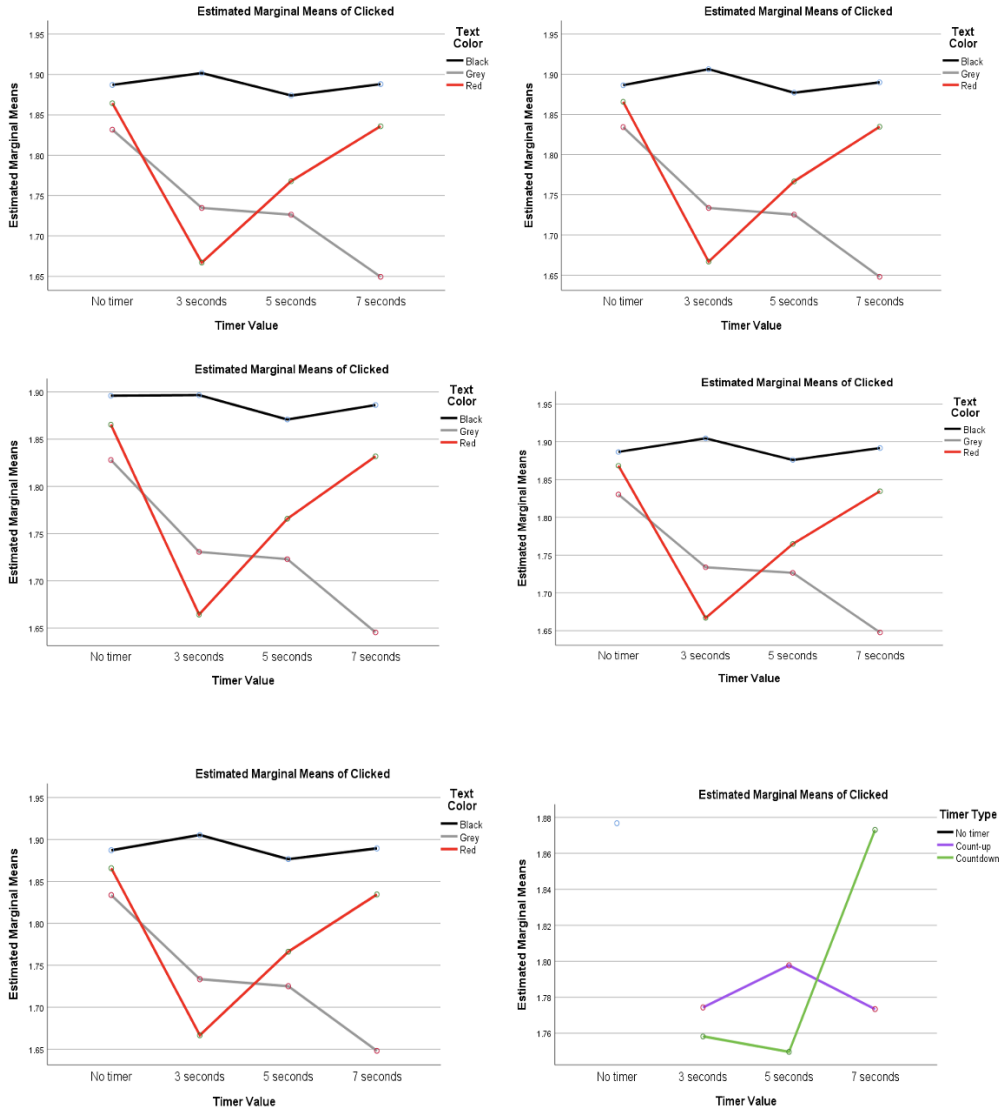

Figure 4. *Profile Plot of Text Color x Timer Value with (a) Age, (b) Gender, (c) Education Level, (d) Volume of Email, and (e) Attention Span as a Covariate, and Profile Plot of Timer Type x Timer Value with (f) Age as a Covariate*

Profile plots of Timer Type x Timer Value with each covariate were performed and appear in Figures 4 and 5. Figure 4(a) shows the profile plot of Timer Type x Timer Value with age as a covariate. Figure 5(a) shows the profile plot of Timer Type x Timer Value with gender as a covariate. Figure 5(b) shows the profile plot of Timer Type x Timer Value with education level as a covariate. Figure 5(c) shows the profile plot of Timer Type x Timer Value with email volume as a covariate. Figure 5(d) shows the profile plot of Timer Type x Timer Value with attention span as a covariate.

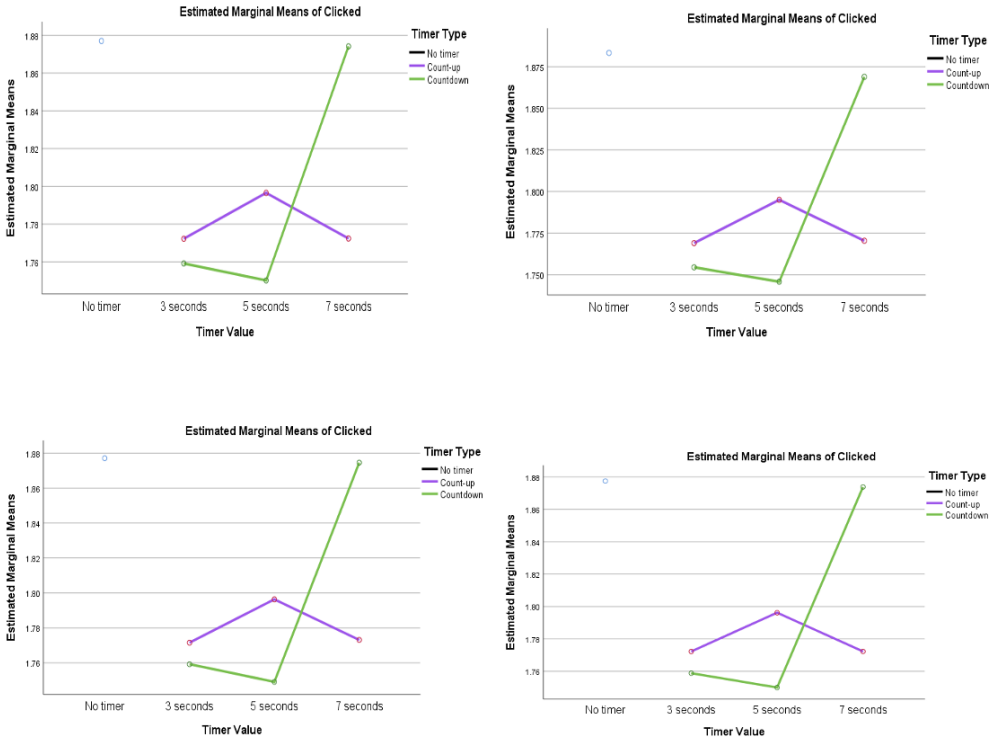

Figure 5. *Profile Plot of Timer Type x Timer Value with (a) Gender, (b) Education Level, (c) Volume of Email, and (d) Attention Span as a Covariate*

Phase III – RQ5 – Age Group

The age demographic was analyzed using the *click* mean and standard deviation of all the email interactions with the simulated phishing emails. A summary of this data is displayed in Figure 6(a). The age demographic that performed the best (had the lowest click mean) was 18-25. The age demographic that performed the worst was Older than 75.

Phase III – RQ5 – Gender Group

The gender demographic was analyzed using the *click* mean and standard deviation of all the email interactions with the simulated phishing emails. The *click* mean for both genders was very similar, indicating that gender may not be a factor in ability to avoid clicking a malicious link or attachment.

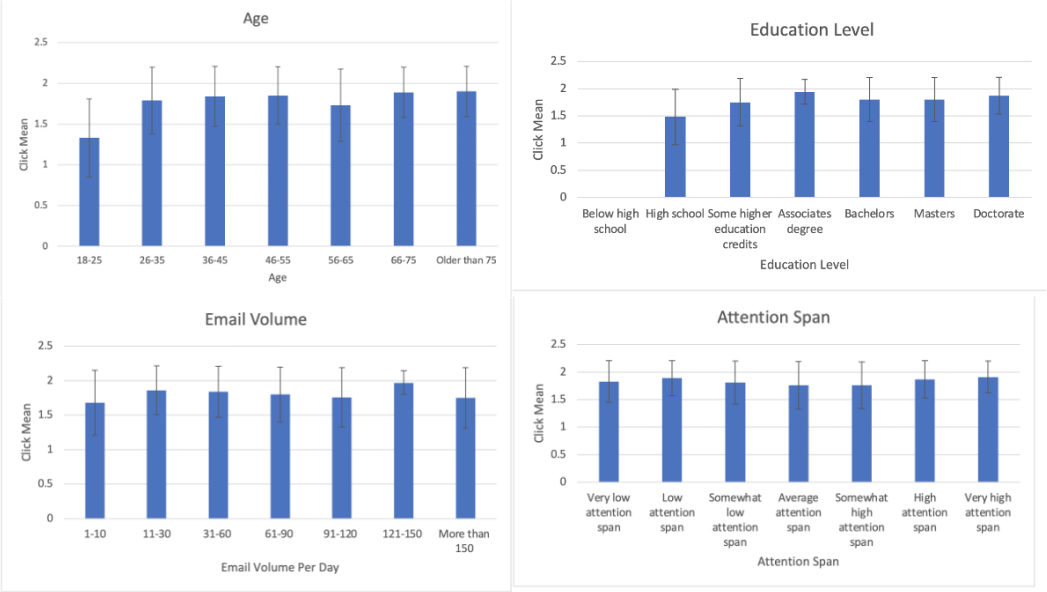

Figure 6. *Summary of (a) Age, (b) Education Level, (c) Volume of Email, and (d) Attention Span Demographics with Respect to Click Mean*

Phase III – RQ5 – Education Level Group

The education level demographic was analyzed using the *click* mean and standard deviation of all the email interactions with the simulated phishing emails. A summary of this data is displayed in Figure 6(b). The Associates degree demographic performed the worst at a click mean of 1.94, and the High school demographic performed the best at 1.48. This indicates that a higher level of education may not mitigate the user's ability to avoid clicking on a malicious link or attachment.

Phase III – RQ5 – Volume of Email Group

The volume of email demographic was analyzed using the *click* mean and standard deviation of all the email interactions with the simulated phishing emails. A summary of this data is displayed in Figure 6(c). The 1-10 emails per day demographic performed the best at a click mean of 1.68, and the 121-150 emails per day demographic performed the worst at a click mean of 1.97. This indicates that fewer emails per day help the user to avoid clicking on a malicious email or attachment.

Phase III – RQ5 – Attention Span Group

The attention span demographic was analyzed using the *click* mean and standard deviation of all the email interactions with the simulated phishing emails. A summary of this data is displayed in Figure 6(d). The Average attention span and Somewhat high attention span demographics performed the best at a click mean of 1.76. The Very high attention span demographic performed the worst at a click mean rate of 1.91. This is counter intuitive as it would be thought that those with a High attention span would be alert to possible phishing attempts.

Summary of Findings

The results and data collection were presented in this chapter. Phase I utilized data from the SME survey to answer RQ1 and RQ2. The PAT mobile app was created and partially tested in Phase II. Pilot testers completed the test of PAT in Phase III. Phase III also included the main study which answered RQs3-5b. An ANOVA was performed on the main study data to answer RQ3 and RQ4. An ANCOVA was performed on the main study data to answer RQ5a and RQ5b. The results of a two-round Delphi process in Phase I indicated values of 3-seconds, 5-seconds, and 7-seconds as the timer values that should be used in the PAT mobile app. Phase I results also validated the sample emails for use in the PAT mobile app as well as the PAT experimental procedure. These data were used in the creation of the PAT mobile app. Phase II resulted in the creation of the PAT mobile app. The app was created using data from Phase I, including the timer values, which sample emails to use, and the experimental procedure. The app was tested using pilot testers. Only minor bugs were found and those were fixed before the main study. Phase III indicated that a countdown timer at 3 seconds with a warning in a text color in red was the most effective in supporting user ability to avoid clicking on a malicious link or attachment. All demographic indicators (age, gender, education level, volume of email per day, and attention span) showed a level of significance.

CONCLUSIONS

Our research demonstrates that red or grey text helps the user's ability to avoid clicking on a malicious link or attachment more than black text does. We also found that a countdown timer is better than a count-up timer or no timer with respect to helping the user to avoid clicking on a malicious link or attachment. Education level appears to have the most positive influence on the user's ability to avoid clicking on a malicious link or attachment both with respect to text warning color and timer value and with respect to timer type and timer value. Younger people seem to have a higher ability to avoid clicking on a malicious link or attachment, and there appears to be no difference in gender regarding the ability to avoid clicking on a malicious link or attachment. It appears that less formal education and receiving fewer emails per day helps one's ability to avoid clicking on a malicious link or attachment. It also appears that a high attention span counters one's ability to avoid clicking on a malicious link or attachment. This is counter intuitive, since it is be expected that individuals with a high attention span would be more likely to have the focus required to analyze a possible phish. The main goal of our research study was to determine through experimental field study whether requiring e-mail users to pause by displaying a colored warning (grey, red, or black text) with a timer (countdown, count-up, or no counter) when they are presented with a potentially malicious link has any effect on the percentage of users falling to phishing attempts. PAT successfully measured user interactions with text warning color and a countdown and count-up timer. The data supports the conclusion that a red or grey warning and a timer, specifically a countdown timer, help the user to avoid clicking on a malicious link or attachment.

Discussion

This study had several limitations. In Phase I, many invalid responses were received, and this is possibly due to the offering of a $10 Amazon gift card. It would have been helpful to ask on the SME survey where they found the survey (Facebook or LinkedIn) as this would have helped to track the source of the invalid data. In Phase II, the outsourced firm for the development of PAT was a limitation. The firm was based in Eastern Europe, so the time zone difference was a limitation in communication. There was also a language barrier that caused some requirements to be misinterpreted which delayed the development timeline. In Phase III, there was a limitation in finding Android users to test the Android version of PAT. A few minor bugs were found, but easily corrected by the outsourced firm. Loading the email simulations into the app was difficult and time consuming. This can be mitigated in future studies by revising the mechanism in which emails are loaded. As it was, each email with each variable value had to be loaded separately, which meant that 21 versions of each email had to be loaded (two timer types (countdown,

count-up) x three colors (black, grey, red) x three timer values (3-seconds, 5-seconds, 7-seconds) + three colors with no timer). During the main study, participants were recruited through Facebook and LinkedIn which caused a limitation of a non-random distribution. Many participants likely had the same characteristics as the researcher. In the first few days of the main study data collection, interaction was low. This was mitigated by posting daily reminders on Facebook and LinkedIn. Also, there were a few minor issues with the simulated emails not showing correctly in the app, but these issues were easily fixed. Many participants were confused by what they were to do despite the directions given. It also appeared that many participants did not read the directions as they asked questions that were answered in the directions. Many participants also stated that they would not have clicked on any of the simulated emails if they had been real. This can be mitigated in future studies by modifying the PAT app to use the user's name as a salutation in the simulated emails.

Implications

There are several implications for cybersecurity and phishing susceptibility reduction. Warning text color and a timer in the warning dialog may play a significant role in user reaction to a possible phish. In addition, age, gender, education level, volume of email received in a day, and attention span may all effect the user's ability to avoid clicking on a malicious link or attachment. While some corporations already present a colored warning dialog to employees when employees receive an external email, *there are no known corporations that employ a timer dialog along with the warning.* Corporations could implement a timer dialog to accompany the existing warning text to provide more mitigation against phishing attacks against their employees. *Our results show that a countdown timer is more effective than a count-up timer or no timer*, lending validation to pedestrian countdown timers. Implications for research indicate that both red and grey warning text may be more effective than black text. Timers have not been used in phishing mitigation research previously, and *these results show that using timers to mitigate phishing is worth further research*. Additionally, our results show that a high attention span negatively effects the ability to avoid clicking on a malicious link which is counter intuitive and that users with a low amount of formal education are more likely to avoid clicking on a malicious link. Future research could investigate these relationships further.

Recommendations and Future Research

The PAT app could be updated to allow for faster loading of email simulations to make it easier to set up a future study. Many participants stated that they would never respond to an email that was not addressed to them. To address this, PAT could be updated to incorporate the user's name in the simulated emails. Multiple participants indicated that they are used to being able to check the actual email address and/or URL by hovering over the presented value. PAT could also be updated to include these features. PAT could also be updated to allow users to categorize emails by junk or valid by assigning the email to a folder and to validate the sender by simulating a block on the sender email. Since the app was coded to auto-populate user simulated inboxes at a particular time of day, the PAT app could be used to explore the effect of time of day on the ability to avoid clicking on a malicious link or attachment. While not used in this study, the warning message is able to be changed in the PAT app, so that a future study could investigate word choice in a warning message. The data collected included whether the participant was using an Apple or Android device although that data was not analyzed in this study. A future study could investigate the effect of device usage on the ability to avoid clicking on a malicious link or attachment including a small device such as a phone vs a larger device such as a tablet.

Summary

In summary, a warning in colored text accompanied by a timer helps users to avoid clicking on a malicious link or attachment. This study indicates that a warning in red text accompanied by a countdown timer is the best combination of text and timer. In addition, this study found that the demographic factors of age, gender, education level, email volume, and attention span all influence the user's ability to avoid clicking on a malicious link or attachment. This study used SME feedback to create a system to investigate whether warning text color or a countdown or count-up timer is effective in helping users to avoid clicking on a malicious link or attachment. The study results showed statistically significant differences among participants presented with red or grey text as compared to black text and presented with a countdown or count-up timer as compared to no timer. Participants were able to notice phishing emails with the assistance of text warning color and a countdown or count-up timer.

References

Amro, B. (2018). Phishing techniques in mobile devices. *Journal of Computer and Communications, 6*, 27-35. https://doi.org/10.4236/jcc.2018.62003

Anderson, B., Kirwan, C., Eargle, D., Jensen, S., & Vance, A. (2015). Neural correlates of gender differences and color in distinguishing security warnings and legitimate websites: A neurosecurity study. *Journal of Cybersecurity, 1*(1), 109-120. https://doi.org/10.1093/cybsec/tyv005

Anderson, B., Vance, A., Kirwan, C., Jenkins, J., & Eargle, D. (2016). From warning to wallpaper: Why the brain habituates to security warnings and what can be done about it. *Journal of Management Information Systems, 33*(3), 713-743. https://doi.org/10.1080/07421222.2016.1243947

Biswas, S., Ghosh, I., & Chandra, S. (2017). Effect of traffic signal countdown timers on pedestrian crossings at signalized intersection. *Transportation in Developing Economies, 3*(1), 2-18. https://doi.org/10.1007/s40890-016-0032-7

Bottazzi, G., Casalicchio, E., Cingolani, D., Marturana, F., & Piu, M. (2015). MP-shield: A framework for phishing detection in mobile devices. *Proceedings of the 2015 IEEE International Conference on Computer and Information Technology; Ubiquitous Computing and Communications; Dependable, Autonomic and Secure Computing; Pervasive Intelligence and Computing*, 1977-1983. https://doi.org/10.1109/CIT/IUCC/DASC/PICOM.2015.293

Bravo-Lillo, C., Cranor, L. F., Downs, J., & Komanduri, S. (2011). Bridging the gap in computer security warnings: A mental model approach. *IEEE Security & Privacy, 9*(2), 18-26. https://doi.org/10.1109/MSP.2010.198

Brooks, T. E., Case, B. J., & Young, M. J. (2003). *Timed versus untimed testing conditions and student performance* I. Pearson Education. http://images.pearsonassessments.com/images/tmrs/tmrs_rg/TimedUntimed.pdf

Brustoloni, J. C., & Villamarín-Salomón, R. (2007, 2007, July 18-20). *Improving security decisions with polymorphic and audited dialogs* [Paper presentation]. 3rd Symposium on Usable Privacy and Security, Pittsburgh, Pennsylvania.

Burns, A., Johnson, M. E., & Caputo, D. D. (2019). Spear phishing in a barrel: Insights from a targeted phishing campaign. *Journal of Organizational Computing and Electronic Commerce, 29*(1), 24-39. https://doi.org/10.1080/10919392.2019.1552745

Butavicius, M., Parsons, K., Pattinson, M., & McCormac, A. (2015). *Breaching the human firewall: Social engineering in phishing and spear-phishing emails*. Proceedings of the Australasian Conference on Information Systems, Adelaide, Australia.

Cheong, L. (2018). *Evaluating visualization for emergency decision-making under uncertainty* [Doctoral thesis, Royal Melbourne Institute of Technology]. RMIT Research Repository. https://researchbank.rmit.edu.au/view/rmit:162600/Cheong.pdf

Chorghe, S. P., & Shekokar, N. (2016). *A survey on anti-phishing techniques in mobile phones*. Proceedings of the 2016 International Conference on Inventive Computation Technologies, Coimbatore, India. https://doi.org/10.1109/INVENTIVE.2016.7824819

FBI. (2018). *Business e-mail comprimise the 12 billion dollar scam*. https://www.ic3.gov/media/2018/180712.aspx

Fine, L. (2016). *The presence of timers and their impact on team communications during high-stress scenarios.* Union College.

Goel, D., & Jain, A. K. (2018). Mobile phishing attacks and defence mechanisms: State of art and open research challenges. *Computers & Security, 73*, 519-544. https://doi.org/10.1016/j.cose.2017.12.006

Goel, S., Williams, K., & Dincelli, E. (2017). Got phished? Internet security and human vulnerability. *Journal of the Association for Information Systems*, *18*(1), 22-44. https://doi.org/10.17705/1jais.00447

Gupta, B. B., Arachchilage, N. A., & Psannis, K. E. (2018). Defending against phishing attacks: Taxonomy of methods, current issues and future directions. *Telecommunication Systems, 67*(2), 247-267.

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. *Social Science Research Network*. https://doi.org/10.2139/ssrn.2544742

Hanus, B., Wu, Y. A., & Parrish, J. (2021). Phish me, phish me not. *Journal of Computer Information Systems*, 1-11. https://doi.org/10.1080/08874417.2020.1858730

Hirshleifer, D., Levi, Y., Lourie, B., & Teoh, S. H. (2019). Decision fatigue and heuristic analyst forecasts. *Journal of Financial Economics, 133*(1), 83-98. https://doi.org/10.1016/j.jfineco.2019.01.005

Hung, L. C., Yang, J. Y., Chen, M. C., Chang, H. L., Ku, C. Y., & Hou, T. W. (2020). Design and evaluation of the bed-cleaning mobile application. *Journal of Nursing Management, 28*(4), 771-776. https://doi.org/0.1111/jonm.12900

Jain, A., Tailang, H., Goswami, H., Dutta, S., Sankhla, M. S., & Kumar, R. (2016). Social engineering: Hacking a human being through technology. *IOSR Journal of Computer Engineering*, *18*(5), 94-100. https://doi.org/10.9790/0661-18050594100

Junger, M., Montoya, L., & Overink, F.-J. (2017). Priming and warnings are not effective to prevent social engineering attacks. *Computers in Human Behavior, 66*, 75-87. https://doi.org/10.1016/j.chb.2016.09.012

Kahneman, D. (2011). *Thinking, fast and slow*. Farrar, Straus and Giroux.

Keegan, O., & O'Mahony, M. (2003). Modifying pedestrian behaviour. *Transportation Research Part A: Policy and Practice, 37*(10), 889-901. https://doi.org/10.1016/S0965-8564(03)00061-2

Klein, G. A. (1993). A recognition-primed decision (RPD) model of rapid decision making. *Decision Making in Action: Models and Methods, 5*(4), 138-147.

Krombholz, K., Hobel, H., Huber, M., & Weippl, E. (2015). Advanced social engineering attacks. *Journal of Information Security and Applications, 22*, 113-122. https://doi.org/10.1016/j.jisa.2014.09.005

Kumaraguru, P. (2009). *Phishguru: A system for educating users about semantic attacks* (Publication No. 3357586) [Doctoral dissertation, Carnegie Mellon University]. ProQuest Dissertations and Theses Global.

Lindahl, C., Wagner, S., Uldbjerg, N., Schlütter, J. M., Bertelsen, O., & Sandager, P. (2019). Effects of context-aware patient guidance on blood pressure self-measurement adherence levels. *Health Informatics Journal, 25*(2), 417-428. https://doi.org/10.1177/1460458217717073

Lo, J. C., Twan, D. C., Karamchedu, S., Lee, X. K., Ong, J. L., Van Rijn, E., Gooley, J. J., & Chee, M. W. (2019). Differential effects of split and continuous sleep on neurobehavioral function and glucose tolerance in sleep-restricted adolescents. *Sleep*, *42*(5), 1-10.

Marto, J. P., Borbinha, C., Calado, S., & Viana-Baptista, M. (2016). The stroke chronometer—A new strategy to reduce door-to-needle time. *Journal of Stroke and Cerebrovascular Diseases, 25*(9), 2305-2307. https://doi.org/10.1016/j.jstrokecerebrovasdis.2016.05.023

Mihelič, A., Jevšček, M., Vrhovec, S., & Bernik, I. (2019). Testing the human backdoor: Organizational response to a phishing campaign. *Journal of Universal Computer Science, 25*(11), 1458-1477.

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

Molinaro, K. A. (2019). *Understanding the phish: Using judgment analysis to evaluate the human judgment of phishing emails* (Publication No. 13424290) [Doctoral dissertation, State University of New York at Buffalo]. ProQuest Dissertations and Theses Global.

Mukhopadhyay, S., & Argles, D. (2011). *An anti-phishing mechanism for single sign-on based on QR-code*. Proceedings of the International Conference on Information Society, London, UK. https://doi.org/10.1109/i-Society18435.2011.5978554

Musuva, P., Chepken, C., & Getao, K. (2019). A naturalistic methodology for assessing susceptibility to social engineering through phishing. *The African Journal of Information Systems, 11*(3), 157-182.

National Eye Institute. (2019). *Types of color blindness*. https://www.nei.nih.gov/learn-about-eye-health/eye-conditions-and-diseases/color-blindness/types-color-blindness

Ndibwile, J. D., Kadobayashi, Y., & Fall, D. (2017). *UnPhishMe: Phishing attack detection by deceptive login simulation through an android mobile app*. Proceedings of the 2017 12th Asia Joint Conference on Information Security, Seoul, South Korea. https://doi.org/10.1109/AsiaJCIS.2017.19

Ndibwile, J. D., Luhanga, E. T., Fall, D., Miyamoto, D., Blanc, G., & Kadobayashi, Y. (2019). An empirical approach to phishing countermeasures through smart glasses and validation agents. *IEEE Access, 7*, 130758-130771. https://doi.org/10.1109/ACCESS.2019.2940669

Newquist, M. H., Dozier, C. L., & Neidert, P. L. (2012). A comparison of the effects of brief rules, a timer, and preferred toys on self control. *Journal of Applied Behavior Analysis, 45*(3), 497-509. https://doi.org/10.1901/jaba.2012.45-497

Oliveira, D., Rocha, H., Yang, H., Ellis, D., Dommaraju, S., Muradoglu, M., Weir, D., Soliman, A., Lin, T., & Ebner, N. (2017). *Dissecting spear phishing emails for older vs young adults: On the interplay of weapons of influence and life domains in predicting susceptibility to phishing*. Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, Denver, CO, United States.

Psychology Today. (n.d.). *Attention span test*. https://www.psychologytoday.com/us/tests/personality/attention-span-test

Ramim, M. M., & Lichvar, B. T. (2014). Eliciting expert panel perspective on effective collaboration in system development projects. *Online Journal of Applied Knowledge Management, 2*(1), 122-136.

Rastenis, J., Ramanauskaitė, S., Janulevičius, J., Čenys, A., Slotkienė, A., & Pakrijauskas, K. (2020). E-mail-based phishing attack taxonomy. *Applied Sciences, 10*(7), 1-15. https://doi.org/10.3390/app10072363

Rosa, E., Dahlstrom, N., Knez, I., Ljung, R., Cameron, M., & Willander, J. (2021). Dynamic decision-making of airline pilots in low-fidelity simulation. *Theoretical Issues in Ergonomics Science, 22*(1), 83-102. https://doi.org/10.1080/1463922X.2020.1758830

Salahdine, F., & Kaabouch, N. (2019). Social engineering attacks: A survey. *Future Internet, 11*(4), 89. https://doi.org/10.3390/fi11040089

Silver, N. C., Drake, K. L., Niaghi, Z. B., Brim, A. C., & Pedraza, O. (2002). The effects of product, signal word, and color on warning labels: Differences in perceived hazard. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 46*(6), 735-739. https://doi.org/doi.org/10.1177/154193120204600611

Sun, J. C.-Y., Kuo, C.-Y., Hou, H.-T., & Lin, Y.-Y. (2017). Exploring learners' sequential behavioral patterns, flow experience, and learning performance in an anti-phishing educational game. *Journal of Educational Technology & Society, 20*(1), 45-60.

Thompson, H. (2012). The human element of information security. *IEEE Security & Privacy, 11*(1), 32-35. https://doi.org/10.1109/MSP.2012.161

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185*(4157), 1124-1131. https://doi.org/10.1126/science.185.4157.1124

Vance, A., Jenkins, J. L., Anderson, B., Bjornn, D. K., & Kirwan, C. B. (2018). Tuning out security warnings: A longitudinal examination of habituation through fMRI, eye tracking, and field experiments. *MIS Quarterly*, *42*(2), 355-380. https://doi.org/10.25300/MISQ/2018/14124

Virvilis, N., Tsalis, N., Mylonas, A., & Gritzalis, D. (2014). *Mobile devices: A phisher's paradise*. Proceedings of the 2014 11th International Conference on Security and Cryptography, Vienna, Austria.

Volkamer, M., Renaud, K., Reinheimer, B., Rack, P., Ghiglieri, M., Mayer, P., Kunz, A., & Gerber, N. (2018). *Developing and evaluating a five minute phishing awareness video*. Proceedings of the International Conference on Trust and Privacy in Digital Business, Regensburg, Germany. https://doi.org/10.1007/978-3-319-98385-1_9

Wogalter, M. S., Conzola, V. C., & Smith-Jackson, T. L. (2002). Research-based guidelines for warning design and evaluation. *Applied Ergonomics*, *33*(3), 219-230. https://doi.org/10.1016/S0003-6870(02)00009-1

Yan, J., Qiao, Y., Yang, J., & Gao, S. (2015). *Mining individual mobile user behavior on location and interests*. Proceedings of the 2015 IEEE International Conference on Data Mining Workshop, Atlantic City, NJ, United States. https://doi.org/10.1109/ICDMW.2015.122