WHO ARE PHISHERS LURING?: A DEMOGRAPHIC ANALYSIS OF THOSE SUSCEPTIBLE TO FAKE EMAILS

Dawn M. Sarno, Joanna E. Lewis, Corey J. Bohil, Mindy K. Shoss, and Mark B. Neider
University of Central Florida

Previous research has identified several populations that are susceptible to inauthentic emails (e.g., spam). However, these studies utilize retrospective, self-report measures to assess email users’ interactions with limited sets of inauthentic emails. In order to fill this gap in the literature, the present study assessed participants’ likelihood to rate a wide variety of emails as spam, authentic, and dangerous. The results highlighted several key findings, 1) there were no gender differences for the email ratings, there were only differences in experience with email, 2) those who do not regularly email and read other electronic documents were more likely to rate emails as spam, possibly indicating an increase in false positives, and 3) the relationship between age and rating an email as spam indicates that younger users may be more susceptible to spam. Overall, the present study identified demographic characteristics that should be considered when training users to detect inauthentic emails.

INTRODUCTION

Cyber-attacks have become more prominent and invasive in recent years. In 2014 Sony Pictures experienced a hack that resulted in the release of more than 47,000 employees’ Social Security Numbers, salary lists and even unreleased films like The Interview (Elkind, 2015). Sony is not the only company to experience a major hack. In 2014 Target also fell victim to a cybersecurity attack when 40 million customers’ credit and debit card information was compromised (Yang & Jakakumar, 2014). These kinds of attacks can often be the result of users clicking a link in a dangerous email. In order to prevent these types of attacks, users need to be trained to be more resilient against inauthentic emails. Before an optimized training program can be developed it is necessary to understand why individuals are vulnerable to dangerous emails. Thus, the present study explored different demographic characteristics and the specific populations that are most susceptible to inauthentic emails.

CYBERSECURITY SYSTEMS AND EMAILS

Emailing is a widespread and necessary form of communication that can threaten many corporate and private cybersecurity systems. This is due to the fact that email has many characteristics that make cybersecurity systems vulnerable to attacks. The difficulty of sustaining attention, discriminability between signal and noise, and failure to follow procedures are all possible vulnerabilities of cybersecurity systems (Boyce et al., 2011). Importantly, all of these vulnerabilities are directly applicable to judging whether or not an email is dangerous. Reading multiple emails requires the user to maintain attention on their task and if this sustained attention lapses, the user may be susceptible to engaging with a dangerous email. Additionally, the signal (inauthentic email) to noise (authentic email) ratio is often very low. Most emails that make their way to a user’s inbox are authentic and safe, which makes users more susceptible when an inauthentic and dangerous email occurs. When these dangerous emails arise users may engage with the email and fall victim to a cyber-attack. This can occur even when the user has specific procedural instructions not to engage with any suspicious emails (i.e., a warning message from the email client or application). Although all users may fall victim to email cyber-attacks, understanding who is the most susceptible is critical in training specific populations to be more resilient.

CHARACTERISTICS OF INAUTHENTIC EMAILS

In order to assess what populations are most at risk to inauthentic emails we have to identify the common characteristics that define inauthentic emails. Dangerous and inauthentic emails are often classified as either spam emails or phishing emails. Many of the characteristics of spam and phishing emails overlap. Characteristics such as using company logos and links, creating a plausible premise, requiring a quick response are all examples of how hackers spoof reputable companies and trick email users into engaging with their emails (Elkind, 2003; Drake, Oliver, & Koontz, 2004; Jakobsson, 2007). Due to these similarities, spam and phishing emails will both
be examined in the present paper and referred to as inauthentic emails.

Wang, Herath, Chen, Vishwanath, and Rao (2012) identified how the types of inauthentic email characteristics influence email users. Specifically, they discovered that the likelihood to respond to an inauthentic email depended on the participants’ attention to visceral triggers and deception indicators. They define visceral triggers as the qualities used by scammers to motivationally manipulate the email user and encourage them to engage with an email (e.g., requiring a quick response). Deception indicators on the other hand are defined as cues to the user that indicate the deceptive nature of the email (e.g., spelling errors). Participants’ attention to these types of triggers was assessed by their interactions with a single email, thus limiting the generalizability to all inauthentic emails. However, the results suggest that certain characteristics of emails are more important for training users to be trained on detecting. Furthermore, these results suggest that some users may be more susceptible to inauthentic emails because of their attention to visceral triggers. Exploring what susceptible populations exist will assist in the development of training paradigms.

SUSCEPTIBLE POPULATIONS

Previous research has indicated that certain populations are more at risk for interacting with inauthentic emails than others (Grimes, Hough, & Signorella, 2007; Sheng, Holbrook, Kumaraguru, Cranor, & Downs, 2010; Kumaraguru et al., 2007; Silva, Emmanuel, McClain, Matzen, & Forsythe, 2015). These studies examine how demographic characteristics such as gender, age, personality, and experience are all involved in a user’s vulnerability to cyber-attacks. It is important to note that most of these studies utilize retrospective, self-report measures, and limited types of emails. However, the identified populations are still useful for future research exploring the relationship between demographic characteristics and engaging with fake emails.

Gender

Research has suggested that women are more susceptible to phishing than men (Sheng et al., 2010). Additionally, other research has found that women rate themselves as lower in computer expertise than men (Grimes et al., 2007). Taken together these findings suggest that women may be more vulnerable to inauthentic emails because they have less confidence with computers and/or are less experienced with computers. However, further work is necessary to understand why women are at more risk for engaging with inauthentic emails.

Age

Older adults. Other demographic characteristics such as age groups have also been explored in their susceptibility to inauthentic emails. Older adults report receiving the same amount of spam despite their lower overall use of the computer (Grimes, Hough, & Signorella, 2007). This indicates older adults may actually receive a higher percentage of inauthentic emails than other populations. Older adults have also been found to be more likely to make a purchase as a result of spam (Grimes et al., 2007). Actually engaging with an inauthentic email and making a purchase highlights the importance of examining how age affects users’ susceptibility to fake emails. It is important to note that these findings were based on self-report data and thus are limited in their conclusions. However, the results indicate that older adults may be more vulnerable to inauthentic emails due in part to a lack of experience and difference in the ratio of fake to real emails. Other age differences have also been found in the engagement with inauthentic emails.

Younger adults. Surprisingly, recent research has indicated that younger adults ages 18-25 are the most vulnerable to email cyber-attacks (Sheng et al., 2010). This finding is surprising in conjunction with the fact that older adults are also found to be more susceptible to engaging with inauthentic emails. Together it suggests that other characteristics related to aging may be involved in the engagement with fake emails (e.g., risk-taking, executive control, experience). Most importantly, many younger adults are extremely active with electronic forms of communication. Electronic communications for younger adults may not be exclusively email, but also other forms like instant messaging and texting. These types of informal communications may closely resemble inauthentic emails due to similar spelling and grammatical errors, possibly encouraging younger adults to interact with them. These findings emphasize the need to train younger adults to be more resilient to email attacks.

Risk Taking

Research has explored how risk taking is related to engaging with inauthentic emails. Interestingly, risk takers have not been found to be more at risk to engaging with inauthentic emails. Instead, users who are more conservative with their risk taking have been found to be more likely to engage with fake emails that were from companies that they do not have an account with
It is suspected that these types of emails are outside of the mental models of conservative risk takers making these types of users vulnerable to new emails. This highlights the importance that future studies should include a wide variety of emails for demographic analysis and training. By including a diverse sample of emails users will be more resilient to all possible emails rather than just the emails they are familiar with.

Novices

Experience has also been explored with users’ interaction with cybersecurity threats. Recently, Silva, Emmanuel, McClain, Matzen, and Forsythe (2015) examined how novice and expert cyber incident reporters engaged with cybersecurity threats. The authors utilized eye tracking and search pattern analysis in an effort to identify to differences in search strategies between novices and experts. Interestingly, novices were found to take longer to find the region of interest and were more easily distracted by erroneous text than expert users. Thus, novices were not only slower than experts to find relevant information, but were less likely to find relevant information. This is directly related to Wang and colleagues (2012) work on visceral triggers and deception indicators. Together these findings suggest that future research should focus on the relevant types of information that are characteristic of inauthentic emails. Additionally, since identification of relevant information in the email appears to be related to experience, expertise should be further explored in the context of vulnerability to in authentic emails.

THE PRESENT STUDY

The previous research is limited in its ability to derive populations who are most at risk to cyber-attacks through email. Most studies either utilized retrospective, self-report measures or have utilized restricted sets of emails. These studies are incomplete because they do not have users actively engaging with emails and users are required to accurately remember their past interactions. Additionally, specific sets of emails are limited in their generalizability to all possible types of inauthentic emails. The present study will fill this gap in the literature by exploring how different populations actively engage with a wide variety of inauthentic emails. Previously identified demographic characteristics such as gender and age will be reassessed with a novel and diverse email set. Additionally, new aspects related to expertise with technology and email will also be assessed in their influence on the likelihood to engage with an inauthentic email. Importantly this study is the first to utilize an email set that specifically includes a diverse sample of inauthentic characteristics (e.g., spelling mistakes, large clickable images, requiring a quick response). Lastly, the present study also takes a novel approach at characterizing the qualities of inauthentic emails, allowing participants to rate emails on the likelihood of being spam, dangerous or authentic. Overall, the present study contributes to literature by assessing previously identified and new populations’ interactions with a novel email set containing a wide variety of inauthentic email characteristics.

METHOD

Participants

A total of 101 participants from the University of Central Florida were recruited for this study in exchange for course credit. The sample was 76% female and ages ranged from 18 to 46 (\(M_{age} = 19.80\)).

Stimuli

Two-hundred inauthentic emails were utilized as stimuli. The emails were either obtained from the researchers’ spam folders or through web searches. A wide variety of types of emails were utilized including banking, advertisements, personal, social media, and shipping confirmations. Typical inauthentic email characteristics were consistent within the email set including using reputable company logos, important links (e.g., verify account), and collecting personal information. The different themes of the emails, as well as the inauthentic characteristics, made the emails more diverse in order to better generalize to all fake emails.

For each email participants were asked to determine the extent to which the email was Spam, Authentic, or Dangerous on a scale of 1-to-7 (e.g. 1, Not Spam to 7, Spam). By asking participants to rate the emails on three different characteristics it provided participants with a greater opportunity to show differences in their interpretations of inauthentic emails. Note that spam was chosen as a characteristic instead of phishing because spam is a more colloquially utilized term.

Procedure

After reading the informed consent all participants completed an online-survey that was created through Qualtrics. The survey consisted of the three ratings (spam, dangerous, authentic) on all 200 emails, a modified Media Multitasking Index (Ophir, Nass, & Wagner, 2009), and demographics. The modified Media Multitasking Index assessed participants use of the
computer for activities such as watching videos, playing games, instant-messaging, e-mailing, reading electronic documents (e.g., pdfs, web pages), and computer applications (e.g., word processing, spreadsheets) (Ophir et al., 2009). For each category, participants were asked if they engaged in that activity and, if so, for how many hours a week. Specifically, the Media Multitasking Index was utilized to assess differences in email and technology experience levels within the participants. The entire survey took approximately 1-hour to complete.

RESULTS

In order to assess the extent to which certain populations may be susceptible to inauthentic emails, participant ratings on the three categories (i.e., spam, authentic, dangerous) were correlated with the modified Media Multitasking Index items and the demographics.

Email Categories

All three rating categories were significantly correlated with one another, spam and authentic, \( r(99) = -.41, p < .001 \), spam and danger, \( r(99) = .49, p < .001 \), authentic and danger, \( r(99) = -.26, p = .010 \) (see Table 1). The positive correlation between danger and spam indicates that as emails seem more like spam they are more likely to be dangerous. The negative correlations between authentic and qualities of danger and spam indicate that the less authentic an email is the more likely that it is both dangerous and spam.

Table 1
Correlations of Email Characteristics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Spam</th>
<th>Authentic</th>
<th>Dangerous</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spam</td>
<td>--</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Authentic</td>
<td>-.41**</td>
<td>--</td>
<td></td>
</tr>
<tr>
<td>Dangerous</td>
<td>.49**</td>
<td>-.26**</td>
<td>--</td>
</tr>
</tbody>
</table>

**\( p < .01 \)

Gender

Unlike previous research gender was not significantly correlated with the likelihood to rate an email as spam, \( r(99) = -.02, p = .856 \), authentic, \( r(99) = -.00, p = .975 \), or dangerous, \( r(99) = .08, p = .437 \). However, gender was significantly correlated with hours of email on the Media Multitasking Index, \( r(99) = .21, p = .043 \), indicating that men spend more time emailing on a weekly basis. Taken together these results suggest that there are no gender differences in the likelihood of rating an email as spam, authentic or dangerous, but rather there are gender differences in the likelihood of experience with email.

Age

Age was not significantly correlated with the likelihood to rate an email as authentic, \( r(99) = .05, p = .63 \), or dangerous, \( r(99) = .11, p = .258 \). However, age was marginally significant with the likelihood to rate an email as spam, \( r(99) = .19, p = .053 \). As the average age was 19.80, it is likely that the relationship between age and rating emails as spam would have been significant with a more diverse sample. As previous research has found this relationship these results suggest that younger adults are less likely to consider emails spam compared to middle age adults. However, given our current sample only contains ages 18 to 46 we cannot make any conclusions about older adults and their vulnerability to fake emails.

Media Multitasking Index

The likelihood to rate an email as spam was negatively correlated with emailing, \( r(99) = -.31, p = .002 \), and reading electronic documents, \( r(99) = -.26, p = .008 \). The likelihood to rate an email as spam was not related to any other dimension of the Media Multitasking Index (all \( p's > .071 \)). Taken together these results suggest that novices are more likely to rate emails as spam than individuals who regularly email and read other electronic documents.

None of the items on the Media Multitasking Index were correlated with the likelihood to rate an email dangerous or authentic (\( p's > .101 \)).

DISCUSSION

Previous research has indicated several populations that are at risk for engaging with inauthentic emails (Grimes et al., 2007; Kumaraguru et al., 2007; Sheng et al., 2010; Silva et al., 2015). However, many of these studies are limited by retrospective, self-report findings, and limited categories of emails. The present study attempted to remedy this gap in the literature by exploring how different demographic characteristics relate to susceptibility of a diverse sample of inauthentic emails.

Notably, the present study did not find a significant difference between men and women on their likelihood to identify an email as spam, authentic or dangerous. However, men were found to spend more time emailing on a weekly basis. Taken together with previous research these findings suggest that women may not be more likely to misidentify emails but are rather less confident
in their expertise with emails (Grimes et al., 2007). It is still unclear if women are actually more vulnerable to falling for inauthentic emails due to this lack of confidence and if training can engender improved performance.

Additionally, other aspects of experience were found to be related to the likelihood of rating an email as spam. Participants’ experience with email and reading other electronic documents (web pages, pdfs, etc.) was negatively related to rating an email as spam. The types of skills necessary to read and assess emails and electronic documents are very similar. Both involve similar aspects like reading blocks of text, url links, and large clickable images. Based on previous research on novices vs. experts it seems that email novices may have been more likely to have false positives due to attention to irrelevant information. Whereas those familiar with email and electronic documents are more likely to assess the email based on learned, relevant characteristics (Silva et al., 2015). Understanding the differences between users with varying levels of expertise will be critical in determining appropriate training methodologies. Specifically, additional work needs to determine if novices make more false positives or if those with experience are more complacent and are actually more susceptible to inauthentic emails.

There was a marginal relationship between age and the likelihood to identify an email as spam. Consistent with previous work, younger adults were less likely to identify an email as spam (Sheng et al., 2010). This suggests that younger adults are more at risk to inauthentic emails than middle aged adults. A more diverse sample may have shown a different relationship between age and susceptibility to rating emails as spam. Specifically, higher susceptibility with younger and older adults and the least susceptibility with middle aged participants. However, the present research is limited due to a limited age range of participants (18-46). In order to disambiguate the differences between age groups identifying spam emails and their interactions with them further research is needed.

Lastly, it is important to note that the present study only found relationships between demographic characteristics and rating an email as spam. No relationships were found between specific populations and rating emails as authentic and dangerous. However, all three characteristics were related to each other. This suggests that there are unique characteristics that indicate to a user that an email is spam that do not also indicate that the email is dangerous and inauthentic. A perfect example of this are advertisements. Advertisements may be authentic and safe to view but are considered spam because they are unwanted.

Multiple classifications are necessary in order to disambiguate how different groups (e.g., older and younger adults) may interpret inauthentic emails. Future work is necessary to assess what specific qualities are consistent across spam, dangerous, and inauthentic emails.

Overall, the present study has identified several demographic characteristics that should be considered when training users to be more resilient to engaging with inauthentic emails. Additionally, since differences in the identification of emails as spam, dangerous and authentic were found, further research is necessary in order to assess what specific qualities of emails lead users to be susceptible to attacks. Ultimately, training paradigms need to be developed in order to train general users and special populations to detect these relevant qualities and be more resilient to email attacks.

**REFERENCES**


