EXPLORING USER STRATEGIES IN DETERMINING TRUSTWORTHINESS OF WEBSITES

by

Manisha Arora

Submitted in partial fulfilment of the requirements
for the degree of Master of Computer Science

at

Dalhousie University
Halifax, Nova Scotia
February 2017

DEDICATION PAGE

I would like to dedicate this thesis to "my husband, Dinesh Tagra who motivated and inspired me to face all the challenges in life and always been on my side whenever I needed him".

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Phishing attacks and breaches in online security are increasing at a high rate, irrespective of current security indicators which aim to warn users against those attacks. We conducted a user study to explore and understand different strategies that users of both technical and non-technical groups follow to determine the legitimacy of websites and emails on their own laptops. We showed websites to all the participants and e-mails to half of them and asked them to determine their legitimacy. This observation session was screen and video recorded. A post-observation questionnaire and semi-structured interview gave us a better understanding of the knowledge and reasons of participants for looking at security cues while making decisions. Based on our results, 67.3% of the phishing websites were correctly identified by our participants on an average (79.2% technical, 55.4% non-technical). While our results were mostly in line with prior research, our use of participants' laptops uncovered a strategy not previously reported. We found that some participants check to see if they are logged in to the website or not to determine its legitimacy, which they can only see while using their own laptops. During our observation, we also identified some differences in the strategies applied by technical and non-technical participants. 50% of our participants who visited websites through e-mails decided about their legitimacy based on the trustworthiness of e-mail. Based on our findings, we provide recommendations that might improve the design of security cues and thus help users in identifying phishing websites more effectively.

LIST OF ABBREVIATIONS USED

SSL	Secure Sockets Layer
TLS	Transport Layer Security
HTTP	Hyper Text Transfer Protocol
HTTPS	Hyper Text Transfer Protocol Secure
EV	Extended Validation
URL	Uniform Resource Locator
T	Technical
NT	Non-technical
Wiki	Wikipedia
Wi-Fi	Wireless Fidelity
CS	Computer Science
EEG	Electroencephalogram

ACKNOWLEDGEMENTS

I would like to convey my deepest gratitude and sincere thanks to my supervisor, Dr. Kirstie Hawkey for supporting, encouraging and guiding me throughout my research project. Her outstanding experience and motivation always inspired me to overcome any challenges that came up during the project. She has been always with me whenever I needed her and made my project smoother to work on. I always admire her and it's always a pleasure to work with her.

I want to give a special thanks to my co-supervisor, Dr. Srinivas Sampalli who taught me the concepts of Security which were very important to form basics of understanding for this project.

I want to thank Dr. Raghav Sampangi for his productive suggestions, which helped me improve the quality of this thesis.

Thanks to Dr. Derek Reilly, for his support during the study and Dr. Bonnie Mackay for a session on Literature review.

I am grateful to Juliano Franz, Ramkumar Velmurugan, Felwah and Fatima who found time from their busy schedules for helping me with my research. Without their help, it would not have been possible.

CHAPTER 1 INTRODUCTION

One of the most important and common challenges of internet security is to protect both the internet users and companies from online phishing scams and prevent them from providing their information to illegitimate websites and emails [7] [2]. Phishing is a "process where a targeted individual is contacted by email or telephone by someone posing as a legitimate institution to lure the individual into providing sensitive information such as banking information, credit card details, and passwords" [1]. It is an attack in which criminals duplicate the company's website in order to gain users' financial or personal information in a fraudulent way [1].

1.1 TYPES OF PHISHING ATTACKS

- **Deceptive Phishing:** This method includes broadcasting email messages to many users with the intention of collecting their personal information by influencing them to click on a fake link. For example: messages stating that account is at risk, or expressing some urgency to change the account details, fill the forms and verify the accounts. Thus, requesting users to enter their confidential information by clicking on the given links [34].
- **Link Manipulation:** In this technique, users receive a link (URL) to a fake website. Clicking on the link opens the fraudulent website instead of the original one mentioned in the link [35].
- **Malware based phishing:** This technique includes installation of malware or malicious software either by clicking on a link in an email attachments or other downloadable malicious files [35].

- **Content-injection phishing:** In this technique, the phisher inserts a fake content on a real website in order to fetch users' personal and confidential information by misleading them [34].
- **Key loggers:** These are a kind of a malware used by phishers to trace the keyboard input. Data that is logged is sent to the phisher, who further decodes and misuses the information [34, 35].
- **Search engine phishing:** In this technique, phishers try to fool internet users by creating fraudulent websites which look legitimate, with eye-catching offers related to various products and services and thus deceive them to give up their details [34, 35].

With an improvement in technology, there has been a tremendous increase in phishing scams. According to the anti-phishing working group (APWG), a total of 289,371 unique phishing websites (figure 1) and 557,964 unique phishing e-mails were reported in the 1st quarter of 2016. There was a prominent increase of approximately 130,000 phishing reports given to Anti-phishing working group over a period of three months as shown in figure 2 [36].

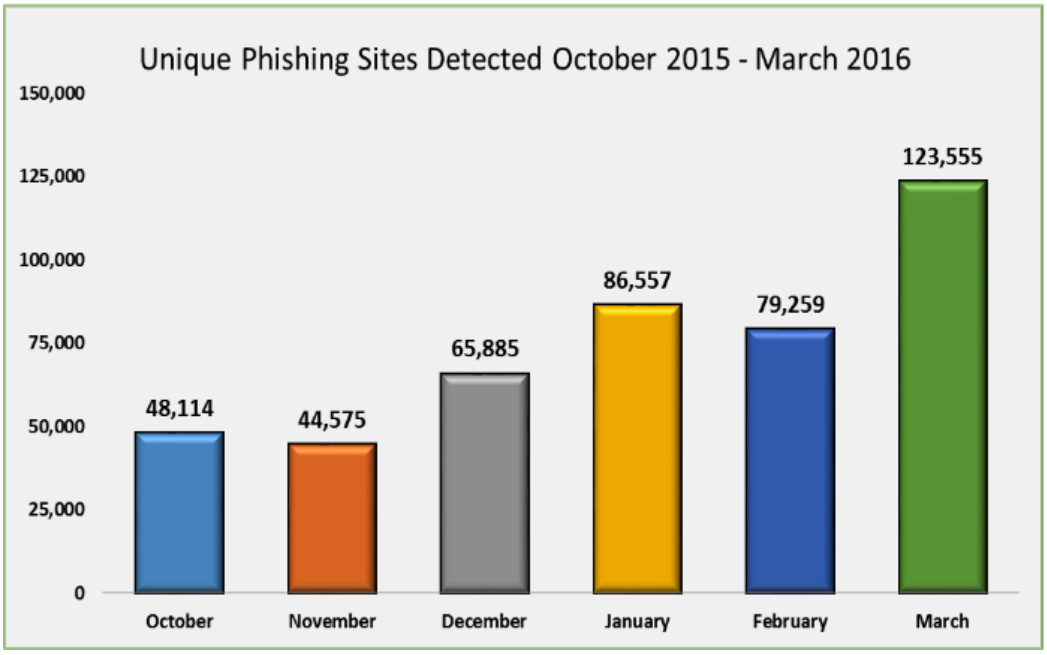

Figure 1 Unique Phishing websites identified from October 2015 to March 2016 [36]

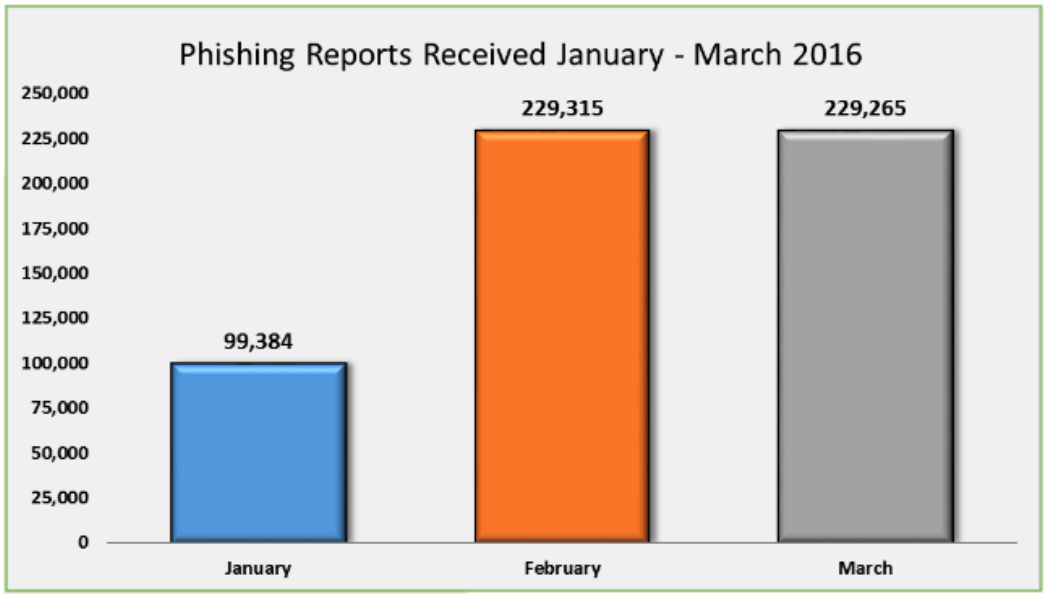

Figure 2 Reports on phishing received by APWG from January to March (2016) [36]

The retail/service sector with 42.71 % of attacks remained the most-targeted sectors followed by Financial services with 18.67% of attacks in the 4th quarter of 2015 [36].

To help users identify phishing websites and emails, security indicators such as the SSL padlock, domain-highlighting, etc. have been designed for web-browsers, but these have not been completely successful in preventing users' from phishing attacks [7].

Many research studies have been conducted in the past to understand users' online behavior and the strategies they follow, which make them susceptible to phishing. Recent studies (e.g., [6, 7, 8, 9, 17]) show that users do not pay much attention to browser security cues. However, most of the research studies are conducted on lab computers, rather than observing users' behavior on their own computers.

1.2 GAPS IN EXISTING RESEARCH

We see the following limitations to this prior work:

1. In the research conducted by Dhamija et al., 2006 [6] and Alsharnouby et al., 2015 [7], the main focus was on non-technical participants. So, by recruiting the technical participants as a different group, we explored the differences in the strategies followed by them as compared to non-technical participants.
2. Dhamija et al., 2006 [6] and Alsharnouby et al., 2015 [7], had participants look at the websites alone, yet phishing emails are frequently the method of attack. Downs et al., 2006 [8] have analyzed the decision strategies of non-expert users (same as non-technical) by showing them suspicious emails through role-play (by giving them fake identities). So, in our case, users have two points of trustworthiness to access: that of the email and then, following the link to go to the website, that of the website. We examined the whole process of users visiting websites through e-mails as well.

3. Use of lab computers: Using lab computer is safe as users' need not provide their own data; however, they may also not be as vigilant as when using their own computer due to this lack of risk. Furthermore, they have to use unfamiliar tools and settings, which may either change their usual behavior or cause them to pay extra attention to unfamiliar warnings and messages. So, we need to find how this artificial environment may impact the results.

Therefore, the purpose of our research is to explore the strategies used by both technical and non-technical users to assess the trustworthiness of websites while accessing websites from their own laptop using browsers of their choice.

1.3 CONTRIBUTION

Our study provided the following contributions:

- We used a mixed design and explored the differences in strategies followed by the technical and non-technical participants, as they determine the illegitimate websites and phishing attacks. For this, we conducted an observation session and semi-structured interviews. We found that there were some strategies which were only followed by some participants in the technical group (such as hovering over the links to check URL's generated by them, checking meta data in inspect element, etc.). We also found that the performance of technical participants in judging the websites correctly (mean: 16) was better than that of non-technical participants (mean: 12.6).

- We explored the differences in decisions made by the participants for judging the websites they reached upon by clicking the link in e-mails with those assessed through website alone. We found that half of the participants who saw websites through emails, judged the website's legitimacy based on that of the e-mail (i.e. if they found e-mail to be suspicious then they were also suspicious about the website linked to that e-mail).
- In our study, participants used their own laptops so, we were able to examine if that makes them more alert and attentive towards the browser security cues and warnings (chapter 4 and 5). We found that participants identified 74.7% of websites correctly, which was almost similar to the findings of prior research (i.e., [6 and 7]) where lab computer was used for the observation session. This finding indicates that participants were not extra attentive or alert towards security cues while using their own laptops.
- We identified that the choice of browser being made during the study has an impact on the findings. We found that the appearance and location of security indicators might influence participant's decision about the website, as compared to the previous studies (i.e., [6, 7, 27]), where researchers have asked the participants to use a particular browser.
- We also explored whether or not participants pay attention to security cues for making decisions about the legitimacy of the websites, and any reasons for the same by exploring further details in semi-structured interview.

- Based on the data collected during the study we provide guidelines to improve the security indicators in the web browsers in order to help users to identify illegitimate websites and phishing emails in order to protect their security online.

1.4 OVERVIEW OF THE THESIS

This thesis is divided in to 7 chapters, which are described as follows:

Chapter 1, provides a brief introduction to phishing, gaps in existing research and our contribution in this area.

Chapter 2, Background and related work discusses the research done so far to analyze user behavior with respect to online security behavior and existing gaps in literature. This chapter concludes with a summary highlighting the gaps in the previous research.

Chapter 3, Methodology discusses the research objective, research questions, and research approach in detail. It then describes the study protocol, instruments used, participant recruitment, data collection, and analysis.

Chapter 4, Results and discussion provides the details of the results obtained from the study. It includes the decisions made by participants on websites and e-mails, strategies used by them in determining both websites and e-mails, and the results of the questionnaire and semi-structured interview.

Chapter 5, Discussion and Implications for design discusses our findings in detail and provides a guideline for a list of suggestions that can be reported for the improvement of web browser security cues.

Chapter 6, Limitations and Challenges discusses the limitations of our study and challenges faced during the research.

Chapter 7, Conclusion and Future work gives the overview of the key findings as a conclusion and the future work that could be done on the basis of the current research.

CHAPTER 2 Background and Literature Review

In this chapter, we present a brief overview of the previous work done in various fields influencing this thesis. Section 2.1 gives the description of the studies done with the perspective of examining the usability and effectiveness of browser security cues in protecting the users against phishing attacks. Section 2.2 describes the prior work done in exploring and understanding user behavior and vulnerability to phishing. Section 2.3 throws a light on the research involving phishing emails. In Section 2.4, research involving the comparison of experts and non-experts has been mentioned along with any findings on the expertise. Finally, Section 2.5 concludes with a summary emphasizing on the gaps in this area of research.

2.1 Effectiveness of Browser Security Cues

Whalen and Inkpen [3], conducted a study on effectiveness of browser security cues. They recruited 16 participants with at least 5 years of web experience and nine out of them were of a technical background. They conducted their study in two phases: the first one being the normal browsing and in the second phase, the participants were asked to pay attention to security explicitly. They used an eye tracker and questionnaire to collect data about participant's interaction with security cues. According to their findings, participants did not pay attention to security cues during phase one whereas in the second phase, they reported some observations. Their results revealed that the lock icon was a commonly noticed security indicator (68.75%) but lacked interaction. Almost half of the participants also checked Https, certificate information was the least noticed security

indicator; even participants with a technical background were unable to interpret it properly. Also, participants did not notice security cues once they signed in to a website. Their findings suggested areas for design improvements in support of web security including improvement in lock symbol, standardizing locations of security indicators in all the browsers, etc.

Herzberg, A. [4] discusses the security indicators, their inability to provide adequate security and possible approaches for improving defense. He gave an overview of various studies related to phishing and browser security indicators. He mentioned that several studies in the past, which focused on the ability of users to detect illegitimate sites, presented alarming results due to high spoof rates. He also mentioned that participants not being exposed to real-life risk and their unfamiliarity to new indicators are the main challenges for the studies. He also discussed various approaches to deal with web security problems such as “identifying bookmarks”, “using password managers”, etc.

Felt et al. [5] conducted a survey with 1,329 participants to evaluate the current security indicators of google chrome. They basically focused on HTTP and HTTPS to gather people’s perception on those. According to their survey results, HTTPS indicator was understood better than HTTP by the participant’s. They found that even tech-savvy participants did not hold correct knowledge about the HTTP indicator of chrome. Due to these shortcomings, they again surveyed participants to evaluate 40 icons and 7 complementary strings created by them. As a result of their analysis they proposed three

indicators (i.e., using “secure”, and “not secure” in URL bar) which have currently been implemented by Chrome.

Egelmen [10] evaluated trust indicators including browser’s phishing and SSL warnings. He created and validated guidelines to improve these indicators. For this, he conducted a study to evaluate the usefulness of phishing warnings by sending 2 phishing emails redirecting participants to phishing websites. As a result, participants were highly vulnerable with 97% trusting the email and proceeding with the link. Active warnings were able to protect users to a greater extent. Recommendations for design improvement included the use of active warnings over passive, which could interfere with user’s task thus diverting their attention to read and follow the warnings. He also focused on changing the look and feel of the phishing website while it is being displayed to the users, to prevent them from trusting it.

Akhawe and Felt [18] conducted a field study to assess the effectiveness of browser security warnings. They used chrome and Firefox’s telemetry framework to collect data and thus analyzed the user’s clicking rates through the warnings. For Firefox SSL warning, the rate of clicking through the warning was only 33%, which reveals that browser’s security warnings can be effective. But click through rates for Chrome’s warning were high (i.e., 70.2%), they found that participant’s experience with warning played a significant role in this user behavior. Thus they found that familiarity and demographic features (such as different operating systems and browser releases) can make a difference in user behavior on how they react to browser warnings. Based on

these, they recommended more work to be done for improving security warnings specifically with relation to demographic factors.

Shi et al. [20] scrutinized the efficiency of web browsers security cues in aiding the end user to correctly identify fraudulent websites. On the basis of previous research they found that the Extended Validation (EV) certificate needs to be improved in terms of design to make it more actionable to users. They adopted affordance based principles and proposed design principles for the EV certificate in Mozilla Firefox. Affordance principle as per them emphasizes on the communication between users and tools [20]. In terms of interface, it indicates that just by seeing the interface end user should be able to perceive its functionality, which could be done by representing it as an icon or symbol [21]. They proposed new design from the perspective of color, design of lock and click label and conducted a user study to evaluate those designs. For the study, they recruited 15 undergraduate university students and asked them to install their proposed designs on their computers in Firefox as add-ons. Later they were asked to respond to an open ended questionnaire concerning the issues and effectiveness of current EV certificates and main features of the new version proposed by them. As a result, they recommended that security indicators should be noticeable, easily understandable and attractive without being annoying.

Wu et al. [23] divided the toolbars into three categories on the basis of normal information (such as domain name, country, etc.), “SSL-verification” toolbar (sites with and without SSL) and “system decision” toolbar (red warning generated by the browser).

During their pilot study, they gave a printed toolbar tutorial to 5 participants and provided it to other 6 through a link in toolbar. They found that the presence of a printed tutorial helped participants to detect attacks, whereas none of the participant in the other group clicked on that link for the tutorial making them fall for attacks. On this basis, they included the tutorial in the main study with 30 participants to evaluate their effectiveness to prevent against phishing attacks.

In this main study [23], they created a sample account "John Smith" on various sites and asked participants to role play as "John Smith's" assistant and to handle 20 emails with a clickable link. They found that user did not pay much attention to security cues in the browser, instead relied on the content of web page. Thus the toolbars were inefficient to prevent users against phishing attacks. Their follow up study consisting of 20 participants was for testing pop up alerts, in which 10 participants saw a normal browser with basic security cues and the other 10 saw a browser with a warning box. The results for participants with a warning box showed low spoofing rates (i.e.10%) as compared to the other group with normal browser (40%). They found that tool bars and security cues were not effective in preventing user against phishing attacks and gave recommendations for improving the websites and design of warnings, such as, use of interruptive warnings, proper differentiation of a real website from a fraudulent one, etc.

Furnell et al. [30] conducted a survey with 340 participants with above average awareness of technology to assess their understanding, knowledge and usability of security features present in Internet explorer, Word, Outlook Express, and Windows XP. Their findings revealed an important need to train users and increase their awareness

about security features and dealing with issues they might phase. They also emphasized the need to improve the security in terms of usability.

2.2 User Behavior and characteristics

In this section, we review the work of authors who have focused on understanding user behaviors and characteristics that make them susceptible to phishing.

Friedman et al. [33] in 2002, conducted a study on how users conceptualize web security. Their study included 72 individuals from three different communities (i.e., rural, suburban, and high-technology). They conducted semi-structured interviews, during which they asked participants about secure connections. They also asked them to identify whether a connection is secure or not by showing them screen shots of web browsers connecting to the websites. Their findings suggest that half of the participants were unable to correctly identify the connection as secure or not; and the ones who were able to identify its security, did not do so for correct reasons. They also found out that in contast to rural and suburban communities, participants from the high-technology community were also not always correct about their understanding of web security; but 92% of high-tech participants were able to correctly identify non-secure connections. They also listed the types of evidence used by participants in recognizing the connections as secure or not including https, type of website, information, lock icon, etc. [32] then used this information to improve their Value-sensitive design work [33].

In 2006, Dhamija et al. [6], evaluated hypotheses about why phishing attacks work and how users are misled by them. They conducted a study consisting of 22 (10 Males and 12

Females) participants with non-computer science background. During the study they displayed 20 websites on a MAC laptop and asked each participant to decide if the website is legitimate or fraudulent and the reason why they think so. *According to their findings, the browsers' security indicators were ineffective to protect users against phishing and led them to commit mistakes 40% of the time.* Participants mainly focused on content and URL together (36%); 23% however, relied only on the content of the website to determine legitimacy of website. They were unable to find any correlation between gender, age, education and other demographic factors with the susceptibility of phishing. Based on their results, they suggested the improvement of security cues based on usability.

In 2015, Alsharnouby et al. [7], conducted a user study to examine whether the improved knowledge of phishing and recent improvements made in browser security cues are effective enough to help users identify phishing attacks. They used a similar scenario for the study as used by Dhamija et al., [6] and used a Windows XP desktop to show 24 websites to participants. They recruited 21participants (9 Males and 12 Females) with a non-computer science background. They also collected eye tracking data to determine whether users have improvements in their mental models of phishing. Their results show that the current browsing cues are also not effective enough to protect users against online phishing as participants were able to identify only 53% of the phishing websites successfully. Thus they gave further recommendations for future designs incuding: changes to URL bar, etc.

Downs et al. [8], in 2006, interviewed 20 non experts with little security knowledge to learn the decisions they take when coming across suspicious emails and websites. In their study they collected data regarding users' knowledge about risks related to phishing, the strategies or decisions they make while reading suspicious emails and the attention they pay to phishing cues. For this, they asked participants to do an email and web role-play exercise. They showed 8 emails to them with 7 emails containing link to webpages, four pop-up messages as images, and 3 websites (1 real and 2 phishing). As per their findings, almost all participants noticed "wrong sender address", "lock icon" and "broken images" present on the web page but didn't interpret them correctly. In deciding the legitimacy of emails, many participants relied on the content of the email rather than other security cues. Their data suggest that users respond more correctly to familiar risks rather than unfamiliar ones and lack in knowledge about phishing and detecting phishing websites and emails.

Downs et al. [9], in 2007, followed up their study with a survey of 232 computer users to better understand how users are susceptible to phishing emails. They asked participants to do a role-playing exercise where their main focus was on finding the factors responsible for susceptibility of users towards phishing. According to their findings, participants with a good knowledge about web environments (i.e., they are able to analyze URL's, etc.) had less susceptibility to phishing attacks. They suggested that educating users regarding browser security cues and also improving these cues so that they are more easily interpreted by non-expert users can help them to avoid phishing attacks in the future.

Sheng et al. [39], in their study with 1001 participants which appears to us as an online survey followed up with a roleplay focused on the demographic factors and their vulnerability to phishing. They found that their female participants who lacked technical knowledge fell for 53.1% of phishing as compared to 41% of male participants.

Thurlby et al. [11] conducted a survey to understand how risk and security is perceived by new undergrad students by analyzing their computer related knowledge. They conducted a survey containing 14 questions with 97 first year undergrad students. Their results summarized that the students related to computing courses represented good understanding of risks as compared to non-computing students. They found that students pay more attention to their own privacy and security in social media and less to the impact on the university's computer systems. So, the researchers recommended further work on information security training and also further investigation on risk security perception and student's knowledge about security threats [11].

Fagan and Khan [12] conducted a survey to explore why users follow or do not follow security advice, which includes changing passwords, using two way authentication, updating software's, etc. They divided the participants into two groups. The first group consisting of participants who follow the advice and the second one of those who don't follow it. According to their findings, the groups have different views on the benefits, cost, and risks related to their decisions on several security practices advised by the experts. According to them, the users who do not follow expert's security advice might not be aware of the benefits of following it, or they might not value those benefits as of

high importance. Thus, they suggested a need to understand the actual thinking and experience of users in order to overcome these gaps in their perception.

Kelley and Bertenthal [13] recruited participants to assess their browser security knowledge through an online survey and its compatibility with the decision they make online. A total 173 participants ranging from 18 to 76 years of age with 100 Males and 73 Females took part in the survey. Their findings revealed that participants responded more accurately to legitimate websites as compared to spoofed ones. They mentioned website familiarity as one of the factors that might have contributed to the ignorance of security indicators. Experts were able to detect fraudulent websites better than non-experts but were not good in detecting websites without any security information. They mentioned "domain highlighting" might be a possible reason for this difference as experts use domain highlighting to make a decision about insecure websites while non-experts do not. Participants who answered 80% of the technical questions correctly, also logged in to 59% of insecure websites. Finally they suggested that instead of a lack of education, inconsistency in website conventions is more of a reason for this risky user behavior.

Kelley and Bertenthal [14] asked their 214 participants to visit and decide the legitimacy of 16 websites ("screen captured" with limited functionality) only by using Firefox. If participants considered a site to be secure, they were further asked to login or else to go back. They kept a time penalty for incorrect responses given by the participants and gave them bonus pay for quickly browsing through the websites. This was followed by an online survey collecting participant's security knowledge and demographic information.

They also collected mouse-tracking data to evaluate participants' behavior leading to their decisions. According to their results, participants with high security awareness performed better than those with low awareness in making a decision to login to a website or not. They also looked over more information before making a decision. Participant's knowledge about the websites, experience, domains, and security indicators all together affect their decision of whether or not log into the websites.

Jakobsson et al. [15] captured user behavior on security indicators of both webpages and emails. They exposed participants to locally hosted emails and web pages. They recruited 17 participants (students, faculty and staff) of 18 to 60 years of age, excluding those with a computer science background. They presented 26 stimuli (emails and web pages) to test on padlocks, logos, grammatical and spelling mistakes, domain names, etc., and asked participants to come up with the factors responsible for authentic appearance of web pages and emails. During this procedure, participants were asked to think aloud their views about the stimuli and were asked to rate them for their legitimacy and fraudulency on a 5-point scale. They didn't allow participants to click on hyperlinks, but scroll and mouse over. All this was screen and voice recorded and followed by a short interview. According to their study, the layout of the page was an important factor; participants paid attention to URL's and also made their decisions on the basis of content rather than other legitimate factors. In the case of emails, those requesting passwords were not considered safe and the ones just containing information as safe. Overall, the participants considered emails as more illegitimate than web pages and said that clicking on links given in emails is a more risky activity.

Neupane et al. [22] used Electroencephalogram (EEG) and an eye tracker to capture how users progress with the security tasks along with their performance. They conducted a 3-d study (dimensions: performance, results of EEG, results of eye tracker) of malware warnings and detecting phishing websites. They found that users paid less attention on the URL area as compared to login and logo. They were able to detect phishing websites only 63% of the time. They found that participants were trying their best to accomplish the given task but they might not be completely aware about it. Thus, highlighting the importance of raising awareness and education on phishing attacks. The results of eye tracker and EEG revealed that participants were actually reading warnings and paying attention to them when presented. Furthermore, participant's personal characteristics like attention also had a direct impact on their vulnerability to phishing attacks.

Jagatic et al. [24] launched a real spear phishing attack targeting university students selected on the basis of information available on social networking sites. To one group email was sent from a known sender and to the other from an unknown sender with the university's email address. Email redirected the victims to a phishing site asking them to enter their login credentials. They were able to determine how many users fall for the attack without actually collecting their credentials. As per their results they found that 77% of females fell for the attack, as compared to 65% of males. Students who were younger were more vulnerable, and those belonging to be from technical fields ("computer science, informatics and cognitive") were less susceptible as compared to students from other fields.

Wright and Marett [25] sent phishing email requesting personal information, which was received by 299 undergrad students. They did this field study to recognize the behavioral reasons responsible for increasing the susceptibility to phishing. They found that 32% of the recipients responded to the email with their information. The main factors to help participants fight against phishing attacks were their web experience and security awareness; those having low score on these were more vulnerable to these attacks.

Sobey et al. [26] examined user response to EV SSL indicators by conducting a user study where they compared Firefox's interface with their own proposed one using an eye tracker. They recruited a total of 28 participants out of which two were from a computer science background. As a result of study, the EVSSL indicator present on Firefox's browser was not noticed by even a single participant whereas almost 53% noticed the presence of the indicator on their proposed interface at least once. However, no participant tried to interact with the indicator to see further information. Also, participants looked for security information mainly on the content of the web page rather than in chrome.

2.3 Spear Phishing

In this section, we have focus on the research done related to spear phishing emails and cues associated to them.

Parsons et al. [16] conducted a study in two phases. In the first phase, they researched and gave a list of 13 cues (such as "consistency", "grammar and spelling mistakes", "sender address", "familiarity", "urgency", etc.) differentiating between legitimate and

fraudulent emails. They asked 5 participants with an information security background to see 50 emails (25 phishing + 25 legitimate) as shown by them in 2013 [38], and rate each of them on a 5-point scale on the basis of the 13 listed cues. They found that personalization, consistency, links, and sender were the features most likely to be found in genuine emails; whereas spelling and grammar mistakes were most likely to be found in phishing emails. In the second phase, they determined the appropriateness of cues used by the participants to decide if the email is legitimate or not. For this they re-examined their study 2013 study [38] and found that participant's decision about legitimacy of e-mails was more based on poor indicators (like visual presentation) rather than appropriate ones. Also, a sense of urgency had an influence on their decision (i.e., participants were poor at managing the emails that reflected urgency). Thus they recommended training and education of users [16].

Lotter and Futcher [19] projected a framework (figure 3) as a solution to address the problem of identifying phishing attacks for email users. They emphasized improving user awareness and education for mitigating these attacks. On the basis of a literature review, they listed several characteristics of phishing emails such as mass emails, spelling and grammatical mistakes, unknown sender, email requesting personal information, etc. They built the framework on the basis of these identified phishing email characteristics; it can be used to categorize the email from highly risky to low risk. This framework can also be implemented in email clients in order to provide users with more information on phishing characteristics of a particular email and hence will help in increasing user's security awareness.

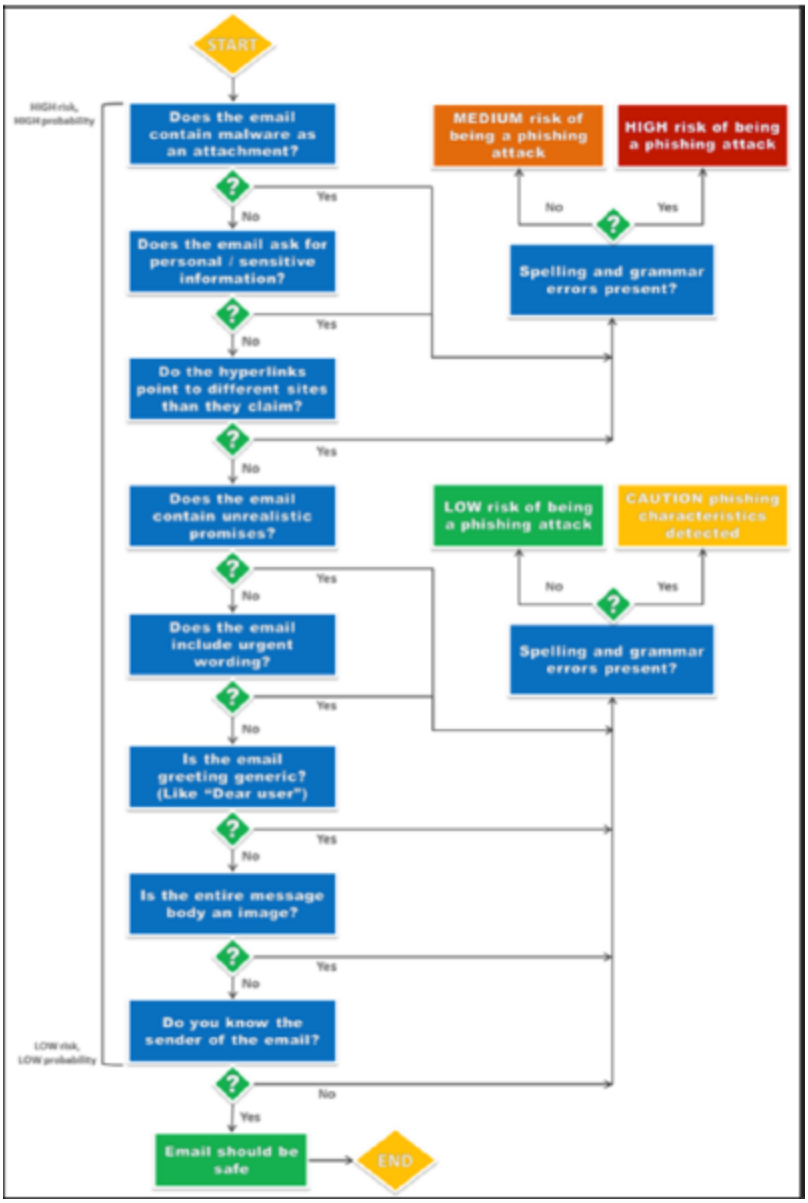

Figure 3 Frame work for identifying phishing attacks in email [19]

Wang et al. [31] conducted a study by using an online survey and the participants for the survey were facers of spear phishing attack. During the study they showed image of the real phishing email to the participants. They asked the participants if they will respond to that e-mail on the likelihood scale. Total 321 surveys were completed with 191 from communication undergrads and 130 from business background. They found that

likelihood to respond e-mails increases if attention is paid to “visceral triggers” like sense of urgency while it decreases if attention is on “phishing deception indicators” such as grammar or spelling error, wrong sender address, etc.

2.4 User Expertise

Recent research by Ion et al. [17], in 2015 examined the online security practices that people follow to stay safe online. They conducted two online surveys with security experts and non-experts respectively. According to their findings, the non-experts users said that to protect their security online they only visit trusted websites. However, this strategy was not often followed by experts as they evaluated trustworthiness of websites dynamically. As a potential solution, the researchers suggested that browsers be developed or improved in such a way that they are able to alert users to detect malicious websites. They also suggested the need to investigate why some non-expert users check or do not check URL indicators, which could help some of the naïve users to avoid phishing attacks and determine website legitimacy.

Kelley et al. [27] conducted a study to analyze the behavior of experts and non- experts in order to discover the causes of failure of security cues. They used eye tracker to find out the behavioral differences between experts and non-experts on their perception of security cues. They asked participants to perform tasks such as: “sharing an item from amazon”, “adding a friend on Facebook from yahoo mail” “logging into yahoo mail using a single sign on”, “commenting on a CNN story” and rating a movie on Rotten Tomatoes” by connecting either through Twitter , OpenID or Facebook [27, 28]. They

asked them to perform these tasks on Dell laptops running Windows Vista using Firefox v4 browser and to use fake accounts given by the researcher to login. This was further followed by a questionnaire asking them about the completion of tasks and a survey for collecting demographic information [27].

The experts were graduate students from Informatics and computing background and were self-identified. They were not able to effectively use the eye tracker, so were unable to get conclusive results for their study. Also, the only criteria they used to differentiate experts from non-experts was the expertise reported by the participants themselves, so they were unable to find differences between both the groups. Thus, they gave recommendations according to the lessons learned during the experiment for other researchers with a similar research objective. Some of their recommendations include: categorizing experts and non-experts on the basis of technical skill questions rather than self -reporting, letting participants select the web browser they want to use, having a diverse study population, etc. [27].

Arianezhad et al. [29] did a similar study as Kelley et al. [27] using eye tracker to determine the impact of technical expertise on the use of browser security cues with perspective of web based single sign-on. They recruited 19 participants including both novices and experts by using survey to categorize them into novice, computer expert and security expert. They found that security experts focused on security cues more than computer experts and novices. Also, all categories of participants had poor understanding of security in web based single sign-on [29].

Results from studies we have previously described also provide findings about expertise.

Friedman et al. [32] compared participants from three communities and found that high-technology were able to perform better than other low-technical communities in identifying secure/non-secure connections correctly. Kelley and Bertenthal [14] showed that participants having high security awareness performed better than those with low knowledge in making decision about logging-in/not logging-in to a website.

Jagatic et al. [24] mentioned that students belonging to technical fields such as computer science or informatics were less likely to fall for phishing as compared to students from other back grounds as shown in Figure 4.

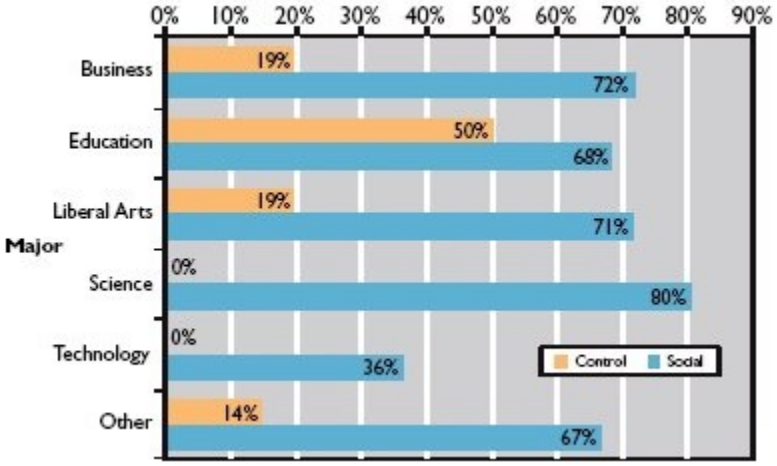

Figure 4 Phishing attack's success rate by participant's education field [24]

2.5 SUMMARY

As mentioned above, prior research describes various challenges, limitations and recommendations to improve the browser security cues and other solutions to prevent users from phishing attacks. However, while understanding and investigating user behavior for susceptibility to phishing, the prior research has had some shortcomings.

For example several studies (i.e., [6, 7, 8]), provided information about the knowledge and strategies followed by the participants, but majority of participants were from non-technical fields so these studies were unable to get insight into the strategies followed by participants from technical fields, specifically with computer science background to decide the legitimacy of websites. Also, these studies were conducted on a lab computer/laptop with a specific browser which would have impacted user behavior.

Some studies (i.e [6, 7]), focused only on websites and didn't explore user response to websites if they visit them through email. Kelly and Bertenthal [14] did a study to understand user behavior in a realistic scenario, but they showed the participants image-mapped websites with limited functionality rather than actual websites with full functionality. Also, their participants were restricted to use only Firefox browser further impacting their behavior and finally, they conducted the study through a survey due to which they were unable to know about the devices used by the participants.

Two studies (i.e., [11, 12]), recommended further investigation of user's perception, knowledge and their actual thinking process. Two studies (i.e., [32, 24]) found that

participants with good technical knowledge performed better than those with less but didn't explore their decision process or strategies which help them perform better. Kelley et al. [27] used an eye tracker but were unable to identify differences between experts and non-experts because of technical limitations of the tool and improper categorization of participants on the basis of expertise. Research done by Ion et al. [17] compared the online security practices followed by experts and non-experts and suggested to determine "why" some non-experts look at URL or not with or without checking Https.

CHAPTER 3 METHODOLOGY

In this chapter, we will discuss the research objective (section 3.1), research questions (section 3.2), study protocol (section 3.3), study instrument (section 3.4), data collection (section 3.5), refinement process (section 3.6), data analysis (section 3.7), recruitment (section 3.8) and participants (section 3.9).

3.1 RESEARCH OBJECTIVE

From the related work, online security breach and phishing attacks are increasing tremendously irrespective of security indicators provided by current web browsers which intend to help protect users against these attacks. These indicators, if enhanced and improved can prove to be of great help for both technical and non-technical online users to protect themselves against those attacks. Through this research, in context of phishing, we wanted to determine the approaches taken by technical users in making decision about websites with which they deal in their daily lives and whether their approach is different from that followed by non-technical users. We wanted to determine how these users respond to the websites when they visit them through email and if there exists any difference in opinion about those websites if they visit them separately. Further, we wanted to know if participants focus on security cues and if not then why not? And to know their perspective about the current cues along with improvements that could be made. Knowing this would give us an insight if they lack in knowledge about those indicators or there exist other reasons to ignore them. Also, we wanted users to use their own laptops with any browser of their choice and determine if that has any impact on

their behavior. From our findings, we wanted to suggest recommendations for designers of the browser's security cues and websites about areas of improvement.

3.2 Research Questions

We have the following high-level research questions:

1. Do participants follow different strategies to determine the trustworthiness of websites, when they visit them directly and via emails?
2. Is there any difference between the knowledge and strategies followed by technical and non-technical participants in detecting legitimate or fraudulent websites and emails?
3. As participants will use their own laptops, will this change their observed behavior and cause them to pay more attention to browser security cues in determining website legitimacy than in prior studies that used lab computers?
4. As compared to the prior research which used a specific browser, does using the browsers which participants use on their daily basis have any impact on the findings?

3.3 Study Outline

Our research consisted of a series of observations, post observation questionnaire, followed by semi-structured interviews, asking participants to assess the trustworthiness of the websites or emails (i.e., they saw whether these are legitimate or fraudulent and to explain the reasons leading them to make this decision). We did this to better understand and explore the strategies used by users to assess the trustworthiness of websites and

e-mails. During the study, participants used their own laptop. A screen recorder software was installed on their machines for data collection purpose and was uninstalled after their session. The questions in the post-observation questionnaire were designed to understand their knowledge and during the semi-structured interview we asked them about their decisions in detail as well as to see if they really follow what they answered in the post observation questionnaire. We recruited 40 participants, including both technical and non-technical with 20 in each category so as to be able to compare the strategies followed by both types.

3.3.1 Study Protocol

After getting approval from Dalhousie's Research Ethics Board (Appendix A), we broadcasted the recruitment script (Appendix C) to various email groups, online classifieds and social media. Participants who showed their interest in participating in the study were sent a link to the screening questionnaire (Appendix E) through which they were assessed for their technical expertise. Participants with low technical expertise were recruited for the non-technical group and those with high expertise were recruited for the technical group of the study.

The study was conducted on individual basis with one participant at a time. At the beginning of the study, researcher explained the whole study to each participant and then the user was required to provide the informed consent (Appendix D).

We used a mixed design in our study and as shown in figure 5, half of the participants of the Technical group accessed half of the websites directly by clicking on links provided on a web page and rest of the websites via emails by clicking on links provided on a web page. The other half participants accessed all the websites directly by clicking on links

provided on a web page. Similarly, half of the participants of the non-technical group accessed half of the websites directly by clicking on links provided on a web page and rest of the websites via emails by clicking on links provided on a web page. The other half participants accessed all the websites directly by clicking on links provided on a web page.

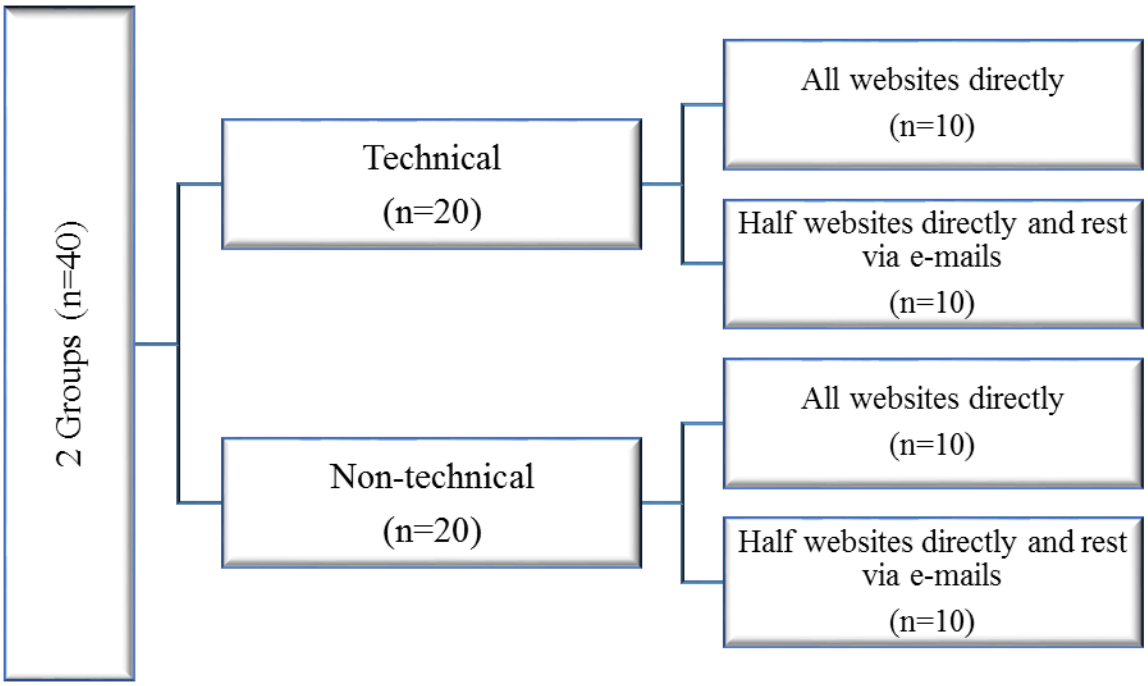

Figure 5 Division of groups

They were asked to judge whether the website or an email was legitimate or not, to explain the reasons leading them to this decision and whether they will login to or proceed further with the website or not. The list of the websites used in this study is adapted from Alsharnouby et al., 2015 [7], with few changes and is shown in Table 1. All illegitimate/phishing websites and emails were created by us and sent from a local server for this study, so that, although the illusion of an insecure website is created, participants were at no actual risk.

During this observation phase, the participants used their own laptops, talked out loud to describe their decision process and screen recording software was installed on their machines to see where they are clicking on the screen, and how much time they spend on each website (it was uninstalled after their session and screen recorded data was collected from their laptop). After testing the set up on our own laptop and doing pilot study, we found that due to technical difficulties an eye tracker software could not be used on participant's laptops, as it takes too long to install and does not collect data properly. So we included the use of video cameras and talk aloud protocol, to analyze where participants are actually looking on the screen while interacting with the websites and e-mails. This gave us detailed information of their decision making process. Video cameras were placed in such a way so that it only captures the participants hand movements and where they were looking on the screen.

This observation session during which websites and e-mails were shown to participants and they were asked to judge them as legitimate or not took about 30 to 35 minutes to complete.

At the end of the observation session, participant completed a Post-observation questionnaire (Appendix F) on Dal Opinio, consisting of questions regarding their knowledge of browser security cues. A portion of this questionnaire was adapted from that used by Ion et al., 2015 [17] in their survey.

Following this, participants took part in a semi-structured interview (Appendix G), where the researcher asked them in detail regarding their decisions about the websites and emails. They were also asked questions related to factors which mainly influenced their decisions to assess the trustworthiness of websites. The interview was audio recorded and

took about 20 minutes. Finally, we compensated and debriefed the participants by telling them the number of emails and websites they correctly identified. We also provided them with information about the security indicators they missed while judging the websites and emails. As an incentive to give correct responses the top 25% of the participants in each category (technical and non-technical), with their scores calculated based on their accuracy in detecting malicious and non-malicious emails/websites, received an additional $10 at the end of the research period. In this way, we hoped that participants will make an effort to give correct answers to what they really think about the websites and emails. The whole session (including the experiment, questionnaire and interview) took approximately 1 hour to complete.

3.3.2 Study Procedure / experimental set up

We created a webpage containing hyperlinks to install screen recording software for both Mac and Windows, websites and websites with emails. Link to websites consisted of hyperlinks to 19 websites which were named as “Website 1”, “Website 2” and so on. Link to websites with emails consisted of hyperlinks to 10 websites and 9 emails, where emails were named as “Email 1”, “Email 2” and so on. We named the links like this so that participants don’t have any clue about the website or email they were to visit. This idea of labeling the websites is similar to the study by Dhamija et al. [6].

All the webpages and websites were hosted from nginx server and could only be accessed by connecting to our lab’s secure Wi-Fi. We bought two domain names from GoDaddy.com and generated SSL certificates to use for those domain names.

We showed 19 websites to participants (7 legitimate and 12 phishing websites). The list of websites and the phishing techniques used in our study were adapted from Alsharnouby et al [7] and is shown in Table 1.

Table 1 List of Phishing and Legitimate Websites Shown during the study

Website	Type	Phishing Technique Used
Bank of Montreal	Phishing	Real website replicated with malicious login link
Netflix	Phishing	Real website replicated with malicious login link
TD	Phishing	Real website replicated with malicious login link, SSL
CIBC	Phishing	Replication of home and login webpages with all other links redirecting to real website
RBC	Phishing	Replication of home and login webpages with all other links redirecting to real website
Scene card	Phishing	Replication of home and login webpages with all other links redirecting to real website
Amazon	Phishing	Real website replicated and overlaid with a malicious pop-up
Twitter	Phishing	Real home page with fake popup asking for username and password in a new window
Dalhousie	Phishing	Misspelled URL, SSL
EBay	Phishing	Real home page with fake popup asking for username and password in a new window
Dalhousie (mydal)	Phishing	Misspelled URL, SSL
Credit card checker	Phishing	Website requesting to enter credit card details
Facebook	Legitimate	Real (SSL)
Kijiji	Legitimate	Real (non-SSL)
LinkedIn	Legitimate	Real (SSL)
TD	Legitimate	Real (EV SSL)
Amazon	Legitimate	Real (SSL)
PayPal	Legitimate	Real (EV SSL)
Netflix	Legitimate	Real (SSL)

We downloaded the original websites directly from the web browser using the save as plugin provided by the browsers and did necessary changes in the code of downloaded websites to use them as phishing websites for our study. We made sure that the website loads properly and all the links present on the first page are redirected to the original website. We tried to do the exact replication of the original website in terms of images and functionality.

Phishing Techniques used:

The techniques we used are similar to those used by Alsharnouby et al [7] and are as follows:

1) Misspelled Domain names: All the phishing websites had either misspelled or suspicious domain names, such as “da1.ca”, “shop-amazon.com”, etc. On clicking any links on the main page of the website, participants were directed to the original pages with changing domain names and certificates.
2) Popups (new window): we created popups which open as a new window and appears on a website when a participant hover over login or sign in button. We used them for ebay and twitter websites.
3) Overlaid popup: This pop up which covers the whole website except URL bar appeared on clicking a button. We used this for Amazon website, to see if participants will be able to relate it to illegitimacy of the website.
4) Context: we designed a page to check the credit card details of the participants. It did not contain any contact information or any other information about the site or organization but an unclickable logo to make it appear more believable.

5) Real websites: We used renowned websites similar to those by Alsharnouby et al. [7] with few changes such as Paypal, TD, etc. By using original TD website we were able to test participants for our phishing website with SSL certificate and original TD website with EV SSL certificate. We couldn't test LinkedIn for http version as done by Alsharnouby et al. [7] because it was getting redirected to the Https version.

We used SSL certificates for few of our phishing websites to show the presence of security on them. We did this because absence of it might have lead participants to detect them as illegitimate easily.

We also showed 9 emails as webpages to half of the participants (8 phishing and 1 legitimate email) containing links to 9 websites. We asked them to imagine that they have received these emails in their inboxes and respond to them as they would normally do in their daily life. We used phishing emails received by us and our friends to create the phishing emails for the study and table 2 shows the list of emails we used.

Table 2 List of e-mails shown during study

Emails	Type	Features
Email 1 (TD)	Phishing	-Asking for account upgradation -Text of link: Upgrade your account here -Not addressed to recipient
Email 2 (Amazon)	Phishing	-Spelling mistake -not addressed to recipient -Text of link: Shop now with Amazon
Email 3 (CIBC)	Phishing	-Account blocked warning -Sign in required -Asking for username and password
Email 4 (Dal)	Phishing	-Warning of discontinuing account: sense of urgency -Misspelled domain name in sender's address -Link: Update your account here

Emails	Type	Features
Email 5 (Paypal)	Phishing leading to real site	-Asking to recover access to account immediately by clicking on a link -Link: recover account
Email 6 (Netflix)	Phishing	-Urgency to verify payment details -Misspelled domain name in senders address -Not addressed to recipient -Link: Verify
Email 7 (Netflix)	Phishing	-Misspelled domain name in sender's address -Real URL to Netflix
Email 8 (Scene)	Phishing	-Misspelled Sender's name and address -Asking to download an app -Link: Download Mobile app
Email 9 (LinkedIn)	Real	-Link: https://www.linkedin.com/ -Addressed to user

We asked the participants if they will click on the link given in the email or not, and asked them to click on it (irrespective of their decision to click or not) to let us know about the legitimacy of the website on which they reached by clicking on that link.

3.3.3 Study Instruments

We used a questionnaire guide for Screening questionnaire (Appendix E) to test the technical expertise of the participants, a Post – observation questionnaire (Appendix F) to test the participants for their knowledge of browser security cues and an interview guide for semi – structured interview (Appendix G) to get details from them regarding their decision about the websites and a coding sheet (Appendix H) for notetaking during the observation session and interview.

Screening Questionnaire

Questions in the screening questionnaire (Appendix E) were created on the basis of related work. Portions of it were adapted from Egelman, S. (2009) [10] and Ion et al.,

2015 [17]. The questionnaire was administered through Dal Opinio to assess the participants for their technical proficiency. It also included background related questions such as age, field of study, gender of the participant, etc.

Post-observation questionnaire

The questions in the questionnaire were planned on the basis of the research questions (section 3.2) and the literature review. Portions of this questionnaire were adapted from that used by Ion et al., 2015 [17] in their survey and the interview protocol of Alsharnouby et al., 2015 [7]. Creating this questionnaire helped the researcher to know about their participant's basic knowledge about browser security cues and what they actually do in their interaction with websites and emails. For example, if they know the meaning of domain highlighting, use of https, focus on URL bar, clicking on a link they received in an email which requests their personal information, etc.

Semi-structured interview guide

During the semi-structured interview, the main aim was to gain in depth details from participants about their decisions during observation phase, the answers they gave in the post-observation questionnaire and what they actually considered in making their decision about the website. The questions in the interview guide were similar to those in post-observation questionnaire and were created based on the related work (Ion et al., 2015 [17], Alsharnouby et al., 2015 [7], Egelman, S. (2009) [10]). We also included the exploration of why's/why not, so as to get more details from participants and to probe them for the reasons of paying/not paying attention to security cues.

3.3.4 Study Instrument Refinement

We piloted our study with 4 peers to check if the links to the websites and emails are working correctly and the questions in the questionnaire and the interview were suitable. All were computer science students with different technical proficiency and helped us in noticing any major flaws missing in the designed websites and emails. They provided their input in pilot testing the survey through Dal opinio and helped in refining the questions both in post-observation and interview guide. After piloting, we added few questions to the semi-structured interview guide about how participants deal with other phishing scams (i.e. phone) and online security. Phone phishing scams are also increasing tremendously, so we wanted to see if participants were aware about these and what strategies they follow to judge them. We were also able to test our screen recording tool on the participant's laptop and how it performs on both Mac and Windows. Thus, pilot testing helped us in fixing both the technical and future data collection problems as far as possible.

3.3.5 Data Collection

In our study, the qualitative data was collected by means of video recordings, audio recordings, screen recorder, note taking and semi-structured interviews. The quantitative data was also collected through screening and post-observation questionnaire.

Video recording

We video recorded the observation part of the session for each participant (both Windows and Mac). Video recording helped us in tracking all the details such as where participants clicked on the screen, on which links they hover over and their hand movements. Video

recording their interaction with the screen helped us gaining the maximum details we required for our analysis without any loss of data and acute details. The quality of the recordings varied depending on the laptops of the participants, but we were able to collect and transcribe data without any loss of information.

Audio recording

We audio recorded the semi-structured interview of participants. This again helped us in preventing any loss of information for this part of the study. If note taking alone would have been preferred instead of audio recording, this would increase the interview time and the researcher would not be able to probe the participants as effectively as was done with audio recording.

Screen recorder

We used Open Broadcaster Software (OBS studio) for screen recording during the observation phase. Use of this software helped us in clearly and easily locating the links and portion of screen visited by the participants. Screen recorder along with video camera together helped us to collect even very minute details and thus good analysis.

Note taking

We took notes during the observation part of the study, where participants interacted with the website. We noted down their responses about the website and emails that if they think whether these were legitimate or fraudulent along with some reasons for making the particular decisions and based on observation. We also took notes during the interview and noted down important points made by the participants.

3.3.6 Data Analysis

The data we collected consist of both qualitative and quantitative nature using various methods. The observation session was video and screen recorded, which helped us in collecting maximum details as possible of the decision made by the participants. All the recordings (video and audio) were transcribed for analysis. We also analyzed the time taken by each participant in judging the website/emails and total time they spent on each website/email. The interview basically focusing on why's/why not was audio recorded and transcribed, followed by coding, in which we allotted different codes to the reasons mentioned by the participants for each question.

We used Microsoft Excel, SPSS, etc. for analyzing the results quantitatively.

3.3.7 Recruitment

The population for the study consists of Internet users. We recruited participants on the basis of their technical expertise by assessing them through a screening questionnaire (Appendix E). We recruited them by broadcasting recruitment notice (Appendix C) through email groups (such as Notice Digest (Today@Dal)), Computer Science Faculty mailing list (csall@cs.dal.ca), etc.). We also posted the notice to online classifieds (e.g. kijiji), social media sites (such as Facebook, etc.) and by putting print outs of the recruitment notice in libraries, learning commons and different departments of our university.

After getting the replies from the interested participants, a screening questionnaire consisting of demographic questions and technical expertise based questions was sent to them (Appendix E). They were assessed for their technical proficiency and were marked on a scale of 0 (lowest) to 5 (highest). Participants marked 0 and 1 were recruited for

non-technical group and those marked 4 and 5 were recruited for the technical group. Those with a score of 2 and 3 were thanked for their interest, but not invited to participate further. This ensured that we were able to detect any differences between the 2 groups. The scores were generated mainly on the basis of three things: participant's field of study, their ability to help others and their response to the options of question 11 in the screening questionnaire. If the participant's were from computer science background with a tendency to help others and have said "yes" to atleast two options asked in question 11. They were given score in the range of 4 to 5. If they were from other fields with a tendency to seek help from others and didn't say "yes" to any of the options in question 11, then they were assigned score from 0 to 1.

3.3.8 Participants

A total of 57 participants responded to our screening questionnaire out of which we recruited 40 participants for our study. Participants were 18 years or older, with at least 1 year of experience with the internet. The age of the participants ranged from 18 to 68 (mean 27, S.D 8.9). Table 3 shows number of males, females, age of the participants taking them separately in two groups: Technical and Non-technical.

Table 3 Participant's Demographics

	All participants (n=40)	**Technical (n=20)**	**Non-Technical (n=20)**
Age (Mean)	26.01	27.1	34.5
Number of Males	20 (50%)	12 (60%)	8 (40%)
Number of Females	20 (50%)	8 (40%)	12 (60%)

Technical participants consisted of 15 Master's (1 full time, 4 part time, and 10 students), 3 doctoral, 2 under grad (1 part time). Our technical participants were mainly from

computer science background except for one participant from library and information studies.

Non-Technical participants consisted of 10 undergrad (1 working part time), 1 retired, 3 professionals, 6 masters (3 working part time), Non-technical participants were mostly from courses like medical sciences, arts, English literature, dentistry, commerce and three from engineering (industrial, civil). None of the non-technical participants were from a computer science background. Further information is provided in the table 4 and 5 below:

Table 4 Technical Proficiency of our participants

Category	Attributes	Overall (n=40)	Tech (n=20)	Non-Technical (n=20)
Seeking Help	Always help others	12 (30 %)	11 (55%)	1 (5%)
	Sometimes help others	12 (30%)	7 (35%)	5 (25%)
	Does not seek other's help	3 (7.5%)	1 (5%)	2 (10%)
	Sometimes asks for help	11 (27.5%)	1 (5%)	10 (50%)
	Always asks for help	2 (5 %)	0	2 (10%)
Proficiency	Designed a Website	22 (55%)	20 (100%)	2 (10%)
	Registered a Domain name	8 (20%)	8 (40%)	0
	Used SSH	17 (42.5%)	17 (85%)	0
	Configured a Firewall	12 (30%)	12 (60%)	0

Table 4 describes the technical proficiency of our participants. 18/20 technical participants were in the category of helping others whereas 12/20 non-technical in that of asking for help. Also, when considered the four attributes of proficiency, only two of the non-technical participants had designed a website, but it was only layout. In contrast, all of the technical participants had an experience with at least two of the attributes as shown in figure 6.

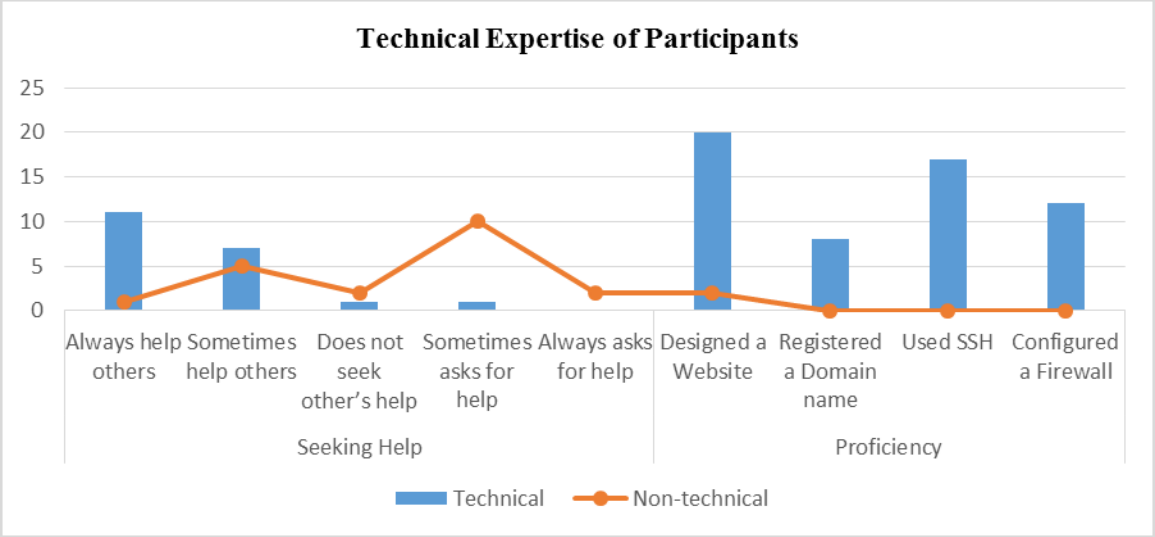

Figure 6 Technical expertise of participants

Table 5 General online practices followed by our participants

Category	Attributes	Overall (n=40)	Tech (n=20)	Non-Technical (n=20)
Use of antivirus		29 (72.5%)	14 (70%)	15 (75%)
Web browser	Google Chrome	31 (77.5%)	16 (80%)	15 (75%)
	Mozilla Firefox	3 (7.5%)	0	3 (15%)
	Safari	6 (15%)	4 (20%)	2 (10%)
	Internet Explorer	1 (2.5%)	0	1 (5%)
Operating System	Mac	16 (40%)	11 (55%)	5 (25%)
	Windows	24 (60%)	9 (45%)	15 (75%)
Online Attack	Credit card fraud	5 (12.5%)	2 (10%)	3 (15%)
	Stolen online password	3 (7.5%)	3 (15%)	0
	Stolen SIN No.	0	0	0
	Identity Theft	3 (7.5%)	3 (15%)	0
Online shopping		37 (92.5%)	20 (100%)	17 (85%)
Check e-mail from other's computer		31 (77.5%)	16 (80%)	15 (75%)

Although if some participants never used a website, we still kept it in our list to gather information on how they judge unfamiliar websites. Table 6 gives the list of common

websites and the number of participants who use them on a particular frequency in their daily lives.

Table 6 List of websites and no. of participants who use them (n=40)

Website	Never used	Use 1-10 times/year	Use 1-10 times/month	Use daily
Amazon.com	3 (7.5%)	21 (52.5%)	15 (37.5%)	1 (2.5%)
Ebay.com	18 (45%)	15 (37.5%)	7 (17.5%)	0
PayPal.com	15 (37.5%)	12 (30%)	8 (20%)	1 (2.5%)
Any Banking website	1 (2.5%)	2 (5%)	26 (65%)	8 (20%)
Social Networking website	0	1 (2.5%)	3 (7.5%)	36 (90%)

CHAPTER 4 RESULTS AND ANALYSIS

In this chapter, we will present the findings obtained from our study. In section 4.1 we discuss about the websites identified as real or phishing by the participants. In section 4.2 we discuss the identified differences based on Technical expertise and Gender. Section 4.3 and 4.4 illustrates the strategies employed by the participants for websites and e-mail respectively. In section 4.5 we mention the responses of participants for visiting some websites through e-mails. Section 4.6 summarize the results of post-observation questionnaire and in section 4.7 we discuss our findings about semi-structured interview.

4.1 DECISION ON WEBSITES

In the observation session, we asked all 40 participants (both technical and non-technical) to visit each website or email one by one and decide whether it was trustworthy or not. We assign the participants a score for correctly identifying both phishing and legitimate websites. This score ranged from minimum 8 to maximum 19 out of 19 websites (Mean 13.9 and SD 3.3). Further, we present the effect of demographic factors such as technical expertise and gender on the results and our findings are as follows:

4.1.1 Technical Expertise

We compared the scores of Technical and Non-technical participants separately to see if there exists any difference. We performed two-way ANOVA to compare both the groups, when taken together a score of correctly identifying the websites we were able to find significant differences with mean score of 16.0 for Technial and 12.6 for Non-technical. This means that technical participants were able to identify the websites correctly more

number of times than were identified by the non-technical participants. The figure 7 below describes how many technical and non-technical participants answered phishing and legitimate websites correctly and table 7 describes the significant differences.

Table 7 Significant score difference of Technical and Non-technical participants based on Accuracy

ANOVA

Dependent Variable: Accuracy

(I) TECH_ID	(J) TECH_ID	Mean Difference (I-J)	Std. Error	Sig.[b]	95% Confidence Interval for Difference[b]	
					Lower Bound	Upper Bound
Tech	Non-Tech	2.938*	.713	.000	1.492	4.383
Non-Tech	Tech	-2.938*	.713	.000	-4.383	-1.492

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

b. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

Figure 7 No. of participants responding to websites correctly based on technical expertise

4.1.2 Gender

Similar to section 4.1.1, when taken score with both phishing and legitimate websites we were able to find significant differences between Males and Females with mean of 15.8 for Males and 12.9 for Females. This means that Males performed better than Females. Figure 8 describes how many Male and Female participants answered phishing website as phishing and non-phishing and table 8 describes the significant differences.

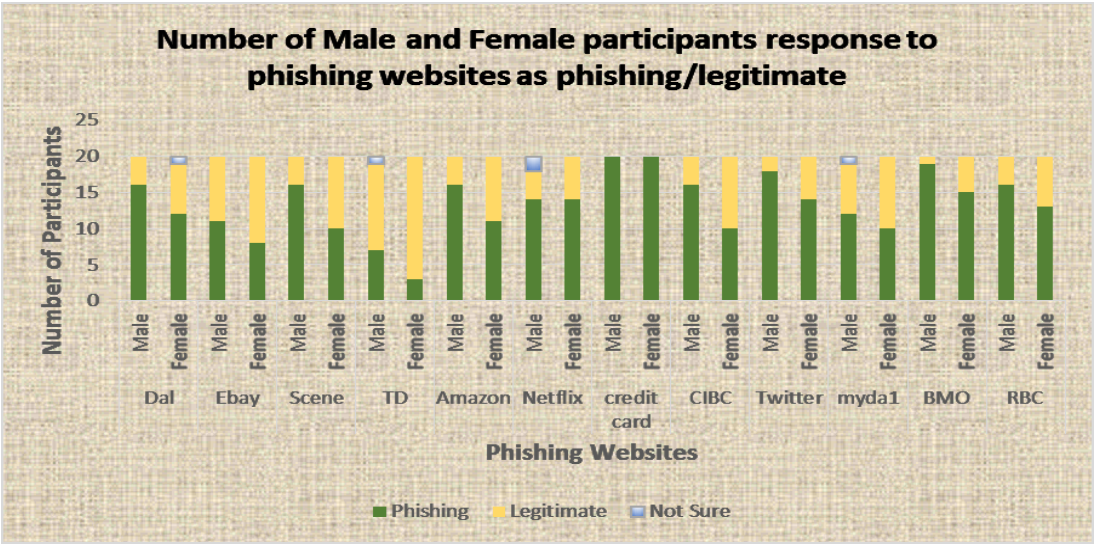

Figure 8 No. of participants responding to phishing websites correctly based on Gender

Table 8 Significant score difference of Males and Females participants based on Accuracy

ANOVA

Dependent Variable: Accuracy

(I) Gender	(J) Gender	Mean Difference (I-J)	Std. Error	Sig.[b]	95% Confidence Interval for Difference[b]	
					Lower Bound	Upper Bound
Male	Female	2.313*	.713	.003	.867	3.758
Female	Male	-2.313*	.713	.003	-3.758	-.867

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

Table 9 Overall Descriptive Statistics of Participants

Dependent Variable: Accuracy

TECH_ID	Gender	Mean	Std. Deviation	N
Tech	Male	16.9167	1.78164	12
	Female	14.6250	2.19984	8
	Total	16.0000	2.22427	20
Non-Tech	Male	14.0000	2.20389	8
	Female	11.6667	2.57023	12
	Total	12.6000	2.64376	20
Total	Male	15.7500	2.40340	20
	Female	12.8500	2.79614	20
	Total	14.3000	2.96302	40

From table 9 which gives the descriptive statistics of Males and Females of both technical and non-technical group, we found that Females of technical group performed little better than the Males of non-technical group. Whereas, the Females of non-technical group didn't performed well as compared to other participants.

Table 10 Between-Subjects Effects

ANOVA

Dependent Variable: Accuracy

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	166.942[a]	3	55.647	11.418	.000	.488
Intercept	7854.704	1	7854.704	1611.604	.000	.978
TECH_ID	82.838	1	82.838	16.996	.000	.321
Gender	51.338	1	51.338	10.533	.003	.226
TECH_ID * Gender	.004	1	.004	.001	.977	.000
Error	175.458	36	4.874			
Total	8522.000	40				
Corrected Total	342.400	39				

a. R Squared = .488 (Adjusted R Squared = .445)

Table 10 describes that there was a significant main effect of technical expertise and gender on the accuracy of judging the websites. And, there was no significant interaction based on the accuracy.

We also compared the scores of dangerous mistakes committed by the participants (i.e. phishing websites answered as legitimate). We were able to find significant differences based on technical expertise and gender as shown in table 11 and table 12 respectively.

Table 11 Significant score difference of Technical and Non-technical participants based on Dangerous mistakes

ANOVA

Dependent Variable: Dangerous

(I) TECH_ID	(J) TECH_ID	Mean Difference (I-J)	Std. Error	Sig.[b]	95% Confidence Interval for Difference[b]	
					Lower Bound	Upper Bound
Tech	Non-Tech	-2.875*	.764	.001	-4.425	-1.325
Non-Tech	Tech	2.875*	.764	.001	1.325	4.425

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

b. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

Table 12 Score difference of Male and Female participants based on Dangerous mistakes

ANOVA

Dependent Variable: Dangerous

(I) Gender	(J) Gender	Mean Difference (I-J)	Std. Error	Sig.[b]	95% Confidence Interval for Difference[b]	
					Lower Bound	Upper Bound
Male	Female	-1.625*	.764	.040	-3.175	-.075
Female	Male	1.625*	.764	.040	.075	3.175

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

b. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

Table 13 describes that there was a significant main effect of technical expertise and gender on the dangerous mistakes committed in judging the websites. And, there was no significant interaction on it.

Table 13 Between-Subjects Effects

ANOVA

Dependent Variable: Dangerous

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	128.167[a]	3	42.722	7.620	.000	.388
Intercept	620.817	1	620.817	110.732	.000	.755
TECH_ID	79.350	1	79.350	14.153	.001	.282
Gender	25.350	1	25.350	4.522	.040	.112
TECH_ID * Gender	.417	1	.417	.074	.787	.002
Error	201.833	36	5.606			
Total	970.000	40				
Corrected Total	330.000	39				

a. R Squared = .388 (Adjusted R Squared = .337)

4.1.3 Successfulness of Websites

In table 14 we summarize the decisions taken by the participants for each website as phishing or legitimate and present the success rate as percentage of participants identifying the website correctly. It also includes, the willingness of participants to login to the website.

The *average success rate* for phishing websites is 67.3% and for legitimate websites 87.5%, which is similar to the results obtained by Alsharnouby et al. [7] (phishing: 53% and legitimate: 79%) with little difference because 20 of our 40 participants were of good technical knowledge.

Table 14 Decision for websites

Website	Actual Type	Answered Legitimate (n=40)	Answered Phishing (n=40)	Willing to Login (n=40)	Success Rate (%)
Dalhousie Uni	Phishing	10(25%)	29 +1(not sure) (72.5%)	7(17.5%)	72.5
Ebay	Phishing	21(52.5%)	19(47.5%)	15(37.5%)	47.5
Scene	Phishing	14 (35%)	26(65%)	10(25%)	65
TD Bank	Phishing	29 + 1(not sure) (72.5%)	11 (27.5%)	19(47.5%)	27.5
Amazon	Phishing	14(35%)	26(65%)	10(25%)	65
Netflix	Phishing	9+2(not sure) (22.5%)	29(72.5%)	7(17.5%)	72.5
Creditcardchecker	Phishing	0	40(100%)	0	100
Cibc	Phishing	14(35%)	26(65%)	8(20%)	65
Twitter	Phishing	8(20%)	32(80%)	8(20%)	80
My.da1.ca	Phishing	17(42.5%)	22+1(not sure) (55%)	14(35%)	55
Bmo	Phishing	6(15%)	34 (85%)	2(5%)	85
Rbc	Phishing	11(27.5%)	29 (72.5%)	9(22.5%)	72.5
Amazon	Legitimate	35 (87.5%)	4+1(not sure) (10	34(85%)	87.5
TD Bank	Legitimate	34 (85%)	6(15%)	33(82.5%)	85
Netflix	Legitimate	36(90%)	3+1 (not sure) (7.5%)	36(90%)	90
Paypal	Legitimate	33 (82.5%)	7 (17.5%)	26(65%)	82.5
Linkedin	Legitimate	37 (92.5%)	3 (7.5%)	34(85%)	92.5
Facebook	Legitimate	37 (92.5%)	3 (7.5%)	37(92.5%)	92.5
Kijiji	Legitimate	33 (82.5%)	7 (17.5%)	31(77.5%)	82.5

Most successful phishing website:

Website with the lowest success rate was a copy of TD bank homepage. Twenty nine participants (72.5%) not correctly identify it as a phishing website. Seven non-technical participants indicated the details present on the webpage were good and correct; 4 non-technical and 2 technical participants thought it had a good layout and that the look of the website appeared to be real; 2 non-technical and 1 technical participants considered it legitimate on the basis of "familiarity". NT19 said "It looks same as I have seen before".

One participant NT6 clicked on a second link leading to login on the original page and had credentials saved. NT6 said "It would be real, it knows my password". One participant judged it on the basis of URL and said "name is correct". One technical participant mentioned the presence of both https and lock as the reason for its legitimacy and 3 participants (2 technical and 1 non-technical) considered only https as the main indicator of its realism. One technical participant (T8) saw that it was not showing a valid certificate when the lock was clicked to see details, then too said it's real. T5 said "I'm a TD user, there are no icons on instructions but it is not asking me to put credit card information". One participant, T16 was not 100% sure about its legitimacy.

27.5% (11 participants) judged it correctly as a phishing website. Two participants (T1 and T5) mentioned the presence of .com instead of .ca in the URL as suspicious. Three participants (T7, NT2 and NT8) mentioned the presence of word "secure" in URL as suspicious and T7 stated that "any company never names it as secure.com". One participant (T14) mentioned the absence of copyrights as the reason for not trusting it. Participant T13 mentioned "there is no encryption". One participant T20 said "I'm not logged in, so it's fake" and participant T6 mentioned the consciousness related to banking website as a reason for not trusting this website. One participant NT9 said "I haven't been to TD before, it's asking me for login but not giving anything to set it up".

4.2 Decision on E-mails

Table 15 summarizes participant's decision about emails and includes the success rate as percentage of participants identifying an email correctly. Emails were seen by total 20 participants (10 technical and 10 non-technical).

Table 15 Decision for e-mails

Email	Actual Type	Answered Legitimate (n=20)	Answered Phishing (n=20)	Success Rate (n=20)	Click on Link (n=20)
TD	Phishing	10 (50%)	10 (50%)	10(50%)	12 (60%)
Amazon	Phishing	5 (25%)	15 (75%)	15 (75%)	6 (30%)
CIBC	Phishing	1 (5%)	19 (95%)	19(95%)	6 (30%)
Dal	Phishing	7 (35%)	13 (65%)	13(65%)	8 (40%)
Paypal	Phishing	6 (30%)	14 (70%)	14(70%)	6 (30%)
Netflix	Phishing	6 (30%)	14 (70%)	14(70%)	7 (35%)
Netflix	Phishing	13 (65%)	7 (35%)	7(35%)	12 (60%)
Scene	Phishing	6 (30%)	14 (70%)	14(70%)	5 (25%)
LinkedIn	Real	10 (50%)	10 (50%)	10(50%)	7 (35%)

We assigned the participants a score for correctly identifying both phishing and legitimate emails. This score ranged from 1 to 9 out of 9 e-mails. (Mean 5.8, SD 1.9).

The *average success rate* for phishing emails is 66.3 %.

4.3 Strategies used by participants In Websites

In this section, we discuss the qualitative representation of all the strategies used by our participants during the observation session to decide the legitimacy of the website. To come up with the strategies we analyzed the screen and video recordings along with the analysis of notes taken on coding sheets (Appendix H). We grouped the strategies into broad ones as shown in Figure 9.

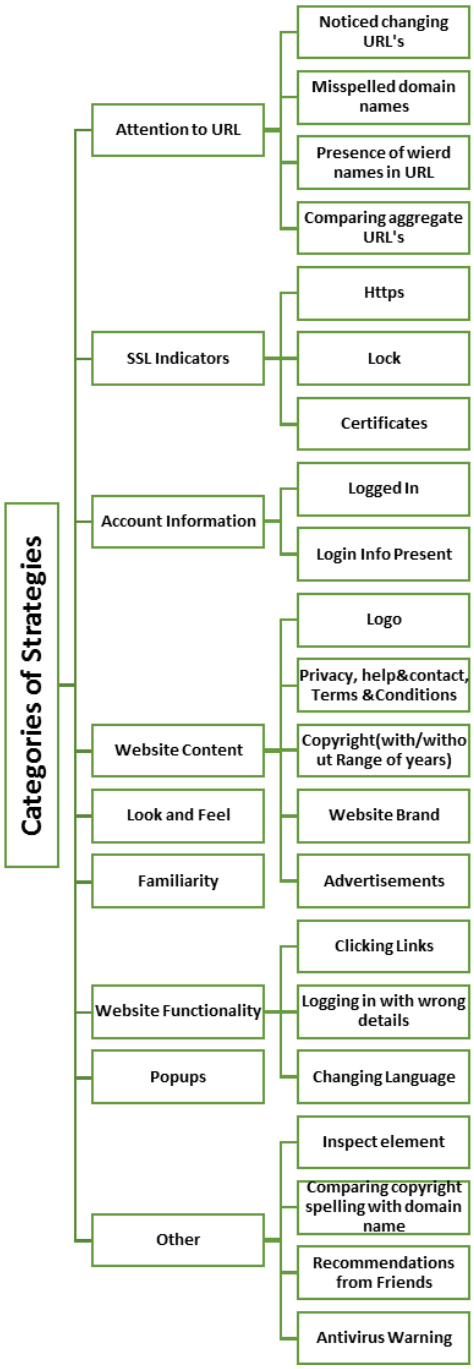

Figure 9 Categorization of strategies

4.3.1 Website Content and Look & Feel

We next discuss the reliance of participants on the content and look & feel of the website to determine its legitimacy. Some of the participants only looked on key content like the presence of a logo; clicking on a logo leading to the home page; the presence of information including "privacy and security", "terms and conditions", "help & contact"; and the history of the website. Almost 12.5% (5) of the total participants solely relied on the content and look & feel without considering any of the browser security cues. All of these participants were from the Non-technical group. Whereas the other 35 participants considered the content along with other strategies to make decisions about the website.

Many (32.5%) of the participants mentioned the absence or presence of logo at least once during the observation session. Three out of 20 technical participants clicked on the logo to see if it takes them to the home page of the website or not. For example, T10 said "I always associate the logo with being a home button". Along with this, one participant (T11) said that "there has to be trademark ™ on top of logo". If it's not there then it's a faulty website.

A quarter of the participants looked for either privacy, terms & conditions, etc. at least once. NT18 said "If I click on the privacy notice, it gives me all kinds of information".

Many (22.5%) checked the copyright information, For example, T3 said "this webpage is somehow new because of 2016, so this provokes a question in my mind whether it's a good website or does it have a good history or not".

One of the participants (T14) said "I don't trust this website, I can't find the copyrights, copyrights are a must on websites, I can't find that anywhere". Similarly, on another website when not finding copyrights, T14 said "I would like to search more like in this

legal tab and all this, whether they have their copyrights somewhere and what are the cookies they are intending to store in my PC".

An interesting finding on the basis of copyright information is that three participants detected a few websites solely on the basis of either years of copyright or information present on it. For example, participant NT14 detected kijiji as phishing and said "just because bottom says copyright 2016 eBay international and unless eBay owns kijiji I didn't know about it, I won't trust it because of whole copyright line". Participant T3 considered if the website is updated to the newest dates on the copyright, it will be real and said "its copyright 2016 eBay international AG, again may be it is a new website, I'm going to think maybe it is legitimate". Whereas, participant NT10 considered the newest dates on the copyright to be phishy and said "copyright 2016, it's really old website, it shouldn't be so soon".

One participant (NT1) revealed "website brand" and five technical participants cited "advertisements" as the judging factors. NT1 stated "It's Amazon, it's a world recognized name, it's not the site that I use, but it does look real". T6 stated on the real TD website that "this message will let me not login because of the word of cookies and the reason they wrote there, it's not convincing enough for me to trust even from beginning". T10 stated "advertisement items are TD products", T18 stated "this website contains few ads like in the sides no it is not real".

Most of the participants also referred to a good design and layout of the website to be one of the reasons for their decision. Participant NT7 only considered the look and feel of the website to make a decision for most of the websites and stated "doesn't look appropriate", "looks authentic"or "format of page looks real". NT11 stated "that's how

kijiji website usually is", "More like professional it's laid down in a nicer way". NT16 mentioned on various websites that "it looks how it is supposed to be".

Two participants (T6 and NT9) raised suspicion if websites on itself asked them to sign in/signup more prominently or the design of website is such that they feel it is urging them to sign in/sign up. NT9 stated "what if I am already a user, I should have an option to cross it out right, like if I just want to sign in and I don't want to sign up then what do I have to do". T6 said "someone wants me to sign in, this looks suspicious for me".

This strategy of relying solely on the content and layout was mostly followed by non-technical participants with or without other factors as compared to few (three) technical participants who solely relied on security cues. So, those participants who didn't see the URL and thought website is similar to what they have previously seen or it looks good in terms of design and organization can be easily fooled by a phishing attack and thus are more susceptible.

4.3.2 Exploring website functionality

Some participants tried to check the functionality of the website to see if all the features present on it are working properly or not. And if they found something working against their expectation, then it raised a suspicion in their minds for that website. A few participants (4/40 NT4, NT10, T8, T18) either tried changing the language on the website or wanted to see what languages are offered by the language change option.

Fourteen participants (8 Technical and 6 non-technical) clicked on various links and advertisements on the website. Some of them clicked to see if those are working and others wanted to see where are these links taking them to? If a link was unclickable or

clicking on it gave an error and if after clicking a particular link it does not take them to a related or expected page it raised a suspicion towards that website. Participant NT4 searched something randomly in two websites.

Only one participant (T4) tested the login field by typing in false information as username and password to check what comes after that. And participant NT20 typed a postal code on one of the websites. Most of these tested features were similar to those found by Alsharnouby et al. [7].

Although we tried to make the websites exactly similar to the original ones along with full functionality, we missed few features on some of the websites. These included an unclickable logo on one, unclickable icons on amazon and also a link on Netflix leading to a 404 not found error. Participants who mostly relied on the fact that clicking a lick should take them to something related page or a particular icon or link should be clickable raised doubts due to these missing features.

4.3.3 Familiarity with the Website

Fourteen participants (5 Technical and 9 non-technical) (35%) reported that they were familiar with a particular website and thus used this strategy together with other strategies to judge the website.

Some of the participants used familiarity to say that the website is real as they know it. Others tried to look for all the things they knew about the website from their experience; and if they found anything weird or unfamiliar, it raised suspicion.

NT6 and NT10 stated on multiple websites “I have been on this website before so it looks similar”. NT1 said “I use TD this look real enough”, NT18 said “it seems to have all same features as original”. NT20 said “This is one of those things which keeps me

connected. I have the habit of logging 4 to 5 times daily. I'm familiar". T12 said "This is Kijiji, I'm familiar with this I will not sign in immediately, and I will check postings which I know about.

Some of the participants (four) when asked to judge the legitimacy of the my.da1 website, if they didn't see URL, then on the basis of familiarity they made a favourable decision on that website.

NT11 said "this is our university website", T3 said "I will proceed with this one because it is exactly the same as the one that I do my login". NT16 said "Dal alert is from today and this is how it works". NT17 said "I saw the alert this morning on this website".

Three participants raised suspicion about the real Facebook website and all three were not signed into it. NT16 said "the logo is little different, it is set up differently, NT18 said "it looks like the Facebook page but something is missing in the logo". And T18 stated that "connecting people image is missing, signup page is different".

Two of the non-technical participants (NT20 and NT17) cited that they never login to unfamiliar websites as they don't trust those.

From this strategy, we conclude that if participants are very familiar with the website, they sometime tend to ignore the browser security cues and believe the website. Thus, they become vulnerable to phishing attacks. This finding is similar to one of the findings of Kelley and Bertenthal [13]. Also, sometimes participants try to find problems with those familiar websites and if they find a glitch they don't fall for it.

4.3.4 Attention to URL

82.5% (33) participants (20 Technical and 14 Non-Technical) used the strategy of paying attention to the URL to judge the website. Similar to what is mentioned by Alsharnouby

et al. [7], if participants were familiar to the websites, they were also aware about the actual URL name and they tried to compare the address present on the website to what they remembered. But in case of unfamiliar websites, or if they were not sure about the address, then they used additional techniques to assess the domain names.

For example, participant T7 matched the spelling of copyright with the domain name and if they were different then judged the website as phishing. One of the participants mentioned that *short URL's* are good. For example, NT2 stated "it doesn't have a long URL, it's very short one" on the Paypal's website.

Some participants mentioned that characters in domain names should be consistent. For example NT9 stated "All the letters are lower case and O is upper case letter, banks don't do this they are consistent". T19 stated "there is zero in url instead of O".

One participant considered the domain name suspicious in itself, NT19 in one of the websites stated "url, it seems too obvious to me that it can lead to theft".

Participant T2 mentioned "address is weird" on several websites as well as participant T9 also mentioned that "it is a bank website and no bank use hyphen instead they chose to have underscore".

Most of the non-technical participants except NT2 and NT3 missed misspelled domain names on a number of the websites as compared to technical participants, who mostly noticed them on all the websites.

Apart from this, in a few websites we used domain names such as "ebay-secure" and cibc-online". Five participants (T2, T6, T7, NT3, and NT7) mentioned the presence of words like secure and online in domain names as suspicious. T7 cited "No bank will name its website cibc-online.com". T6 said "word secure in domain name is suspicious".

T2 said “I don’t think bank will give this domain name” and “If I create a website I will not use secure in website name. It’s not security”.

In contrast, one non-technical participant (NT 16) considered the presence of secure in the domain name as safe and stated “this is secure so that means it is safe to use”.

The other URL feature that five (T1, T5, NT10, NT16, NT18) of our participants noticed was the presence of .ca/.com at the end of the website address. Participant’s T1 and T5 mentioned the presence of “.com” in our TD website as illegitimate. Also T5 said “In Canada it should be amazon.ca not .com”.

For our Netflix website named as netfiix.com, NT10 said “website address is wrong, it’s not .com it’s .ca” because of familiarity with the URL.

Two other participants (NT16 and NT18) checked the presence of “Netflix Canada” in the website content and then noticed the URL address to be “netfiix.com”, upon which they said that the domain name should be .ca. However, both the participants didn’t recognize that netfiix was misspelled.

Changing URL’s: The other important thing was noticing changing URL’s when participants were redirected to an original page from the phishing page by clicking a link on it, and while coming back to the phishing page from the original one. Six technical participants (T2, T5, T6, T7, T13, and T15) were able to notice this change correctly and four (T3, T8, T10 and T17) couldn’t identify the changing URL’s. Participant T7 clicked several links in a new tab on our Da1.ca website and justified the original and fake URL by comparing the fake page URL with all other redirected original links.

In case of non-technical none of them recognized this change except NT2 who was able to see it but couldn't interpret it correctly. All other participants didn't clicked on link to get redirected to different pages.

Three technical participants (T7, T8 and T9) *hovered over various links on the website and compared the aggregate URL's, which appeared while hovering over the links at the bottom of the page to the domain name of the website*. None of the non-technical participants followed these strategies. T9 mentioned that the host name is the first thing to monitor as soon as website is opened because of the past incident for scotia bank. He stated "they started a new domain name and I was visiting that web page so I was shocked that how can you change that. I just sorted that out by confirming it with their customer care and they told this has been done for a time being. To confirm domain name correctness, T9 stated "even if I hover over this image, check the URL which has been created, and then compare both, so host name is different the things which have been populated are from different domains" also he noticed changing URL's to confirm incorrect domain names.

4.3.5 Attention to SSL Indicators

The majority of participants (65%; 8 non-technical and 19 technical) mentioned the presence/absence of one of the SSL indicators at least once while making decisions about the website. They used these indicators along with other strategies to identify the truthfulness of the website. Seven technical participants (T1, T2, T4, T7, T8, T9, and T12) paid attention to all three https, lock icon and certificate information.

For the other 12 technical participants, three (T13, T14 and T18) paid attention to both lock and https; one (T16) to both lock and certificate information, five (T3, T5, T10, T11, T15) looked only at https, and three (T6, T19 and T20) only at the lock icon. T20 mentioned for one of the websites "green thing secure". In the case of the eight non-technical participants, one (NT2) looked at lock and certificate information (verified by whom) at least once. One (NT4) paid attention to both lock and https at least once, whereas out of remaining 6, NT10 only looked at certificate information by clicking on lock icon at one website but could not interpret it correctly. Two (NT3 and NT15) looked only at https and three (NT5, NT14 and NT19) only looked at lock icon at least once during the observation session. This is summarized in table 16. The reasons why or why not participants look at the SSL indicators are mentioned in Section 4.6.

Table 16 Summary of SSL indicators

NT (n=8)	T (n=19)	https	Lock	certificate
Φ (null)	7	√	√	√
1	3	√	√	-
1	1	-	√	√
2	5	√	-	-
3	3	-	√	-
1	Φ (null)	-	-	√

4.3.6 Popups

Only nine (8 technical and 1 non-technical) participants noticed popups and reasoned it as a phish indicator for the website. T11 said "when you move a cursor on sign in it will not open a page like this". T14 said "I won't proceed because it is giving too many pop-ups, I don't think Twitter is so desperate to say please login". We embedded pop ups on 3 websites; on ebay and twitter it appeared as a new window when the cursor go over the sign in option, and on amazon it appeared on the same window covering the whole

screen, when one clicked on “sign in securely”. So a few participants were not able to see these if they didn’t cross over or click on a particular button during their session. Of the 17 participants who actually did this, only 9 (8 technical and 1 non-technical) were able to notice them and 8 non-technical ignored them. Out of these 5 of the participants had blocked pop-ups and they didn’t pay attention when their browser showed a pop-up blocked signal.

In addition to the strategies mentioned above which were similar to those found by Dhamija et al. [6] and Alsharnouby et al. [7] we also found additional strategies, which included account information, antivirus warnings, as presented next.

4.3.7 Account Information

Just over half (21/40 52.5%) of participants (10 technical and 11 non- technical) identified websites as legitimate if they were logged in to it (i.e. if they found there name and information on it or their username and password were saved by their browser for that website). Two participants NT6 and T20 also mentioned that if they are not logged in to the website on which they have saved their username and password then it’s definitely fake. For example, T20 said “usually whenever I go to dal login it is my browser that gives me option like it’s always logged in.” NT6 said “I wouldn’t say that it’s real because Netflix just automatically logs me in” when she didn’t find her username and password saved. Out of other 19 participants, NT2 (who used our laptop) also mentioned that if not logged in to the website, then it’s a fake one.

This strategy can only be found when participants perform the study on their own laptops; Prior research (i.e. [6, 7]) was unable to find this strategy because participants

used lab computers to judge the websites. This strategy can also prove beneficial for protecting users against phishing attacks and highlights that users who follow this strategy cannot be tricked by those attacks.

4.3.8 Antivirus warning

Three participants (7.5%; NT1, NT15, NT17), saw Web Advisor's warning on one of the websites during the session. On seeing the warning, they judged the website as fraud. All three participants were from the non-technical category. NT1 said "MacAfee advisor, it's saying that it's not a good place to go, so I'll say it's not real". NT15 said "I think there is something wrong with this, because it's not opening properly". This highlights that if participants are aware that their antivirus has a warning about website, then they tend to believe it.

4.3.9 Other strategies

Two participants (T12 and NT4) preferred doing a google search to verify one of the websites. T12 stated "I'll do background check, I'll go to wiki page, they have website link and will compare it". NT4 said "I'll google it to check if LinkedIn goes by this link only or it's something different and then I'll login".

One technical participant, T20, checked *Meta data in inspect element*, opened another tab and typed the URL from Meta data to compare the two websites.

None of the non-technical participants used these techniques to judge the legitimacy of websites. One of the non-technical participant NT19 mentioned recommendation from friends as one of the reasons for showing belief in one website and stated "lot of my friends recommend me for using it".

Figure 10 summarizes the number of technical and non-technical participants who followed a particular strategy.

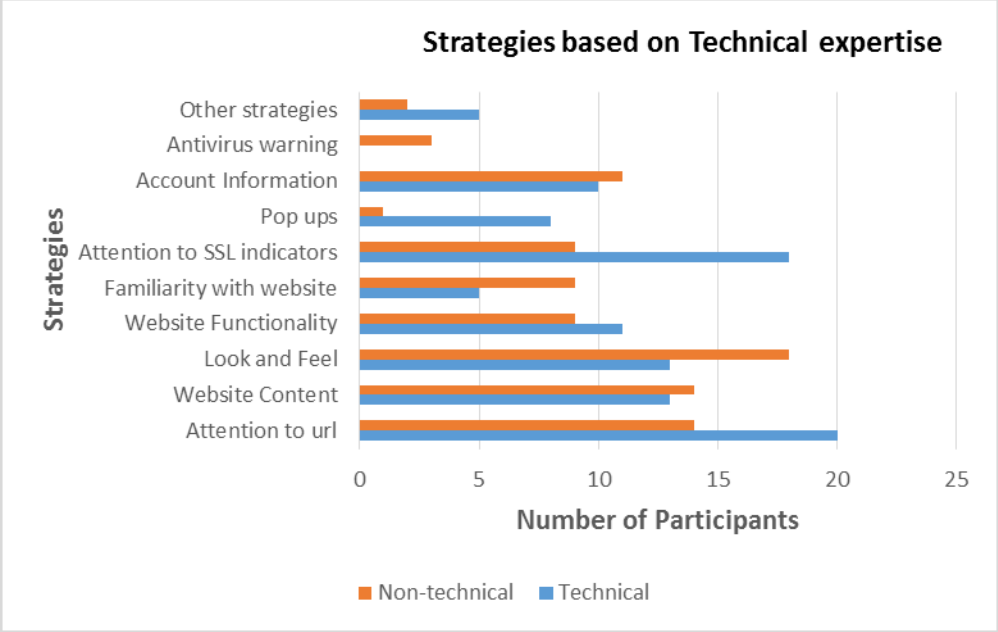

Figure 10 Strategies based on Technical expertise

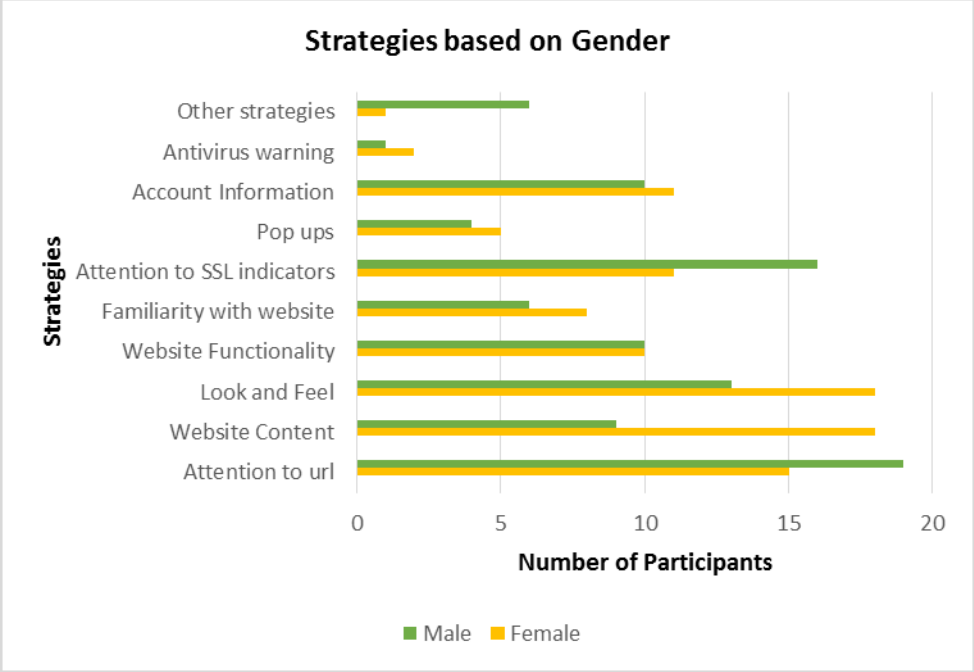

Figure 11 Strategies based on Gender

As shown in figure 11, non-technical females more relied on website content and look and feel as compared to technical females and males.

Four of our participants couldn't use their own laptops, so we tried to make the environment more realistic as possible by providing them a laptop with the same operating system and letting them use the browser of their choice. We didn't noticed any differences in strategies used by them except that they could not make use of account information to judge the websites. So we have provided the results including them, after clearing it in debriefing session that they follow the same strategies on daily basis.

4.4 Strategies Used by participants in Emails

In this section, we provide a qualitative representation of all the e-mail strategies used by the twenty participants who during the observation session decided the legitimacy of the e-mails. To come up with these strategies we transcribed the screen and video recordings along with the analysis of notes taken during the session. Table 17 describes the main strategies identified by both the technical and non-technical participants. The majority of the participants focussed on the sender's email address (16/20) and content of the e-mails to judge them. For example, NT11 said "Microsoft always provide you information to update but doesn't say that you would not be able to reach your account unless you update it", NT14 stated "In no way an email should ask you to reply with your account information, they should know it". T12 stated "this looks like a threat, there is no customer service tone in this email".

Only three participants mentioned importance of subject line in judging emails, T12 said "I always read subject before going to content". NT14 stated in one of the e-mails "my issue with this, why I don't trust it is, it comes under this subject line, it doesn't seem

right and makes me feel that it's not from LinkedIn" Four participants said that there should be name always in salutation instead of member or customer. T20 said "when anything comes from Dal they don't say Dal user, they will call you from your own name". Eight participants were always doubtful about emails related to banks and preferred to call them rather than following link given in the emails. For example, T12 said "I'll prefer to call about anything that has to do with my account, either it is secure or not, usually I call them". NT20 said "I won't reply it like this, I will prefer calling customer service tell them do it online, I will definitely not send my username password here".

Two participants said if there is a link to download something, it should not take you to the homepage of the website but to actual page from where you can download that thing. For example, T13 said "normally when you click on the links like download the mobile apps, it should directly download the apps, instead of taking you to some website." Only five participants paid attention to the link present in the email and four participants clicked the link at least on one email before making decision about it. For example, T19 said "I'll probably click on link and see how official it is".

Table 17 Number of participants as per strategies followed in e-mails

Strategy		**Technical (n=10)**	**Non-technical (n=10)**
Sender's e-mail address		10 (100%)	6 (60%)
Subject of e-mail		2 (20%)	1 (10%)
Salutation		3 (30%)	1 (10%)
Content of e-mail	Spelling errors	5 (50%)	5 (50%)
	Extreme words	3 (30%)	9 (90%)
	Personal information required	6 (60%)	5 (50%)
	Logo/brand	1 (10%)	2 (20%)
Copyright		2 (20%)	4 (40%)
Attention to link address		3 (30%)	2 (20%)
Clicking link leading where		1 (10%)	1 (10%)
Financial e-mails		5 (50%)	3 (30%)
Clicking link before deciding e-mails		3 (30%)	1 (10%)

We considered the number of participants in each category if they followed it at least in one of the e-mails.

4.5 VISITING WEBSITES VIA E-MAILS

In this section, we want to illustrate the decision of participants on websites when they visited them via a link present in e-mail. The actual response of whether they wanted to click on the link or not is mentioned in Section 4.2. After that response, we asked participants to click on links anyway to make a final decision about the websites. After analyzing their decisions we found that four of the ten technical (T17, T12, T19 and T18) and six of the ten non-technical participants (NT11, NT12, NT16, NT18, NT19 and NT20) who saw 9 websites through email stated that they either won't login or won't believe the website as legitimate because they don't trust the email from which they reached to the website.

For example, T17 said “I don’t trust e-mail so phishing”, NT11 said “Netflix real, but won’t proceed due to email, if I open separately by myself then I’ll login”. Also three participants, two technical (T17 and T20) and one non-technical (NT20), decided a few websites were legitimate based on the trustworthy email. For example, T17 said “Netflix real due to authorized email”. NT20 said “Netflix is real because of trustworthy email”.

4.6 Post- Session Questionnaire and Interview

As discussed in Section 3.3.1, at the end of the observation session participants were asked to fill a questionnaire based on their knowledge of security cues and general habits of dealing with e-mails on daily basis, which was further followed by a semi-structured interview. In this section, we discuss the results obtained from questionnaire and interview.

Understanding of https:

We asked the meaning of https in the questionnaire. All 20 technical participants were able to answer this question and we received various responses. So, we categorized them as follows, 17 participants related its meaning to “security”, two participants (T4 and T20) mentioned that “website has a certificate”, and one participant (T7) mentioned “that it’s a protocol”.

Out of 20 non-technical participants, 13 didn’t know what it means, one participant (NT10) said that “site is real”, three participants (NT3, NT4, NT15 and NT19) related it to “security and site safety”, one participant, NT14 said “website is certified”, and two participants (NT2 and NT14) mentioned that it’s a protocol and website hosted is secure.

In the interview, we further probed participants on checking the presence of Https and reasons for whether or not they did it. When asked if the participants checked the presence of Https, all the participants had three main responses "yes", "no" and "for banking and websites requesting credentials" as shown in Figure 12.

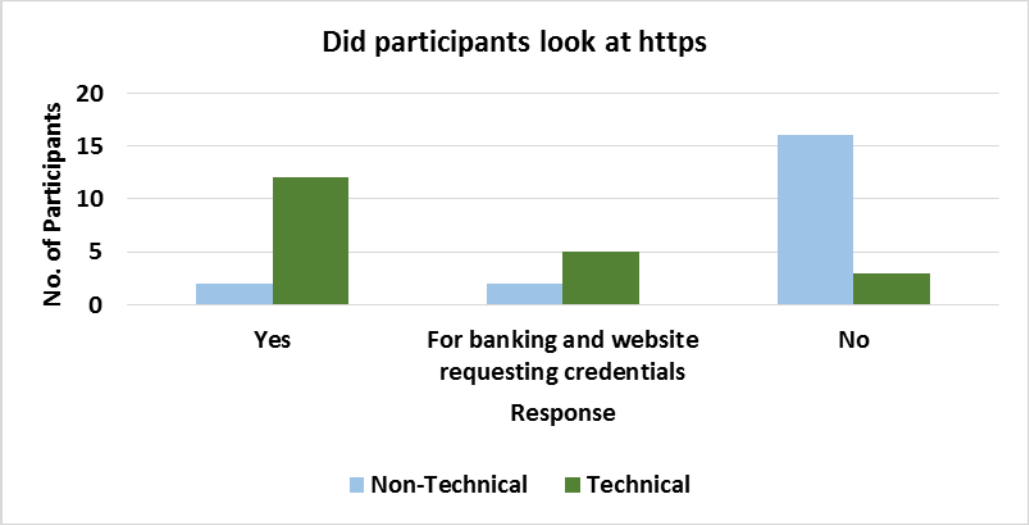

Figure 12 Did participants look at Https

In case of technical participants, T1 didn't check https for the first website and on asking said "Not for first website may be because it was dal and I am using it every day so if you see whole thing you don't see link, it didn't occur to me first to verify." T1 also mentioned that in general "I didn't used to be conscious before but after getting banking accounts I became more conscious".

T5 said "If I'm doing research I might forgot to check because on that time I focus on content not on the site maybe I will not look at https but if I'm logging at a site which is taking my credentials then I'll look at it and make sure it is actual site". T17 and T19 forgot to check during the study but T17 stated that they usually check it. T19 said "I don't know what it means when it doesn't show up". Three participants (T6, T16 and T20) said that they checked its presence but according to our observation they didn't.

For non-technical participants, NT10 didn't mentioned https while visiting websites during study. N14, NT7 and N5 didn't check https during study but usually stated that they check it for banking websites or while logging in. NT7 said "when I login I check it, but not for websites here".

When asked for reasons why they check it, main reason given by the technical participants was "Security" as can be seen in figure 12. T6 said "if it is green it makes me feel real", T9 said " the last letter 's' is the only thing which states it's a secure thing, even if there is no lock, if there is https it's a secure connection made for the website".

One participant, (T11) had an incident in the past and started checking https after that. T11 said "once I had got stuck because of fake Microsoft website, they asked me to call them on a toll free number, they took control of my entire computer by asking me to put username and password, which they provided".

Non-Technical participants gave various reasons, we categorized them as can be seen in figure 13, which included trusting antivirus. As NT2 and said "I do trust my antivirus, I have gone to sites in the past where they haven't been safe and I have been warned. I don't see how much harm they could do". Three participants (NT3, NT5 and NT10) checked it for website security. NT10 said "I do not look it to verify website as sometimes real websites does not have Https, may be it doesn't specifically signify that a website is real". Three participants (NT4, NT15 and NT19) mentioned that they studied about it in school. NT15 said "I did a session in school about managing accounts online, so I know about it". To five of the participant's (NT9, NT14, NT16, NT18 and N20) it didn't strike them to check Https. NT14 said "It didn't catch my attention here". The

remaining 8 participants (NT1, NT6, NT7, NT8, NT11, NT12, NT13 and NT17) were not aware about it so they didn't check its presence.

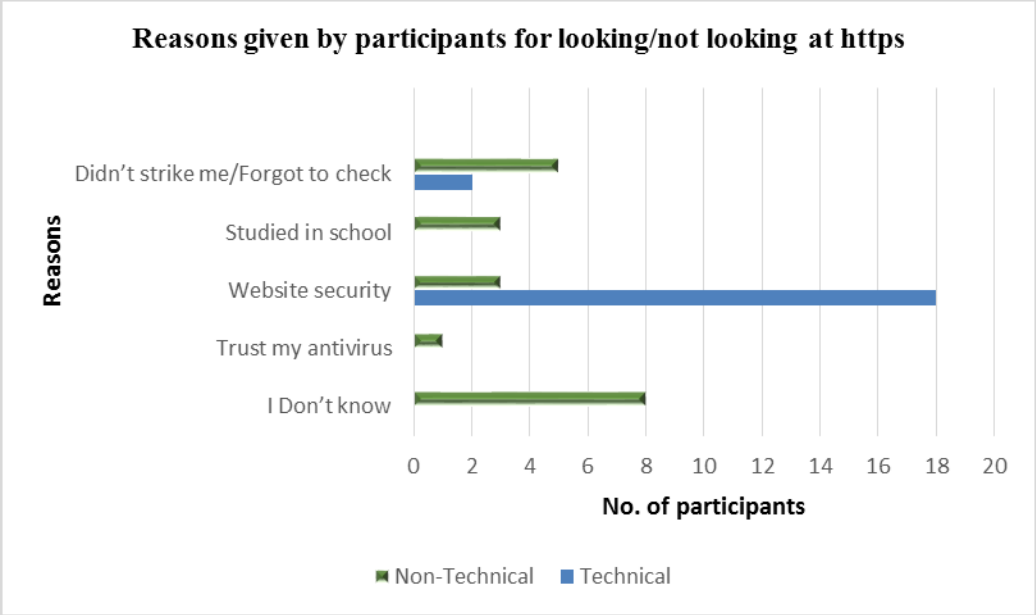

Figure 13 Reasons given by participants for looking/not looking at Https

Understanding of Lock Icon:

We asked participants meaning of the lock icon in our post-study questionnaire. One of the technical participant (T5) didn't know what it indicates, other 19 technical participants mentioned that it's related to website security and connection is verified.

Five non-technical participants (NT1, NT4, NT7, NT9 and NT18) didn't knew what it indicates. For the other 15 participants who did, ten related it to site security, one participant (NT6) said "browser is secure", and participant NT20 said "restricted/blocked".

In the interview, we probed them for checking the lock icons presence and the reasons for doing it or not. The majority of technical participants (17/20) said that they looked at lock

icon whereas 13/20 non-technical did not as can be seen in figure 14. Three technical looked at it most of the time (T2, T8 and T18) and three didn't looked (T5, T17 and T10). One technical participant T19 looked only when it showed up. Four technical (T10, T13, T14 and T20) and five non-technical (NT2, NT5, NT7, NT14 and NT19) looked at it only for banking websites or other websites requesting credentials. Three technical participants (T3, T11 and T15) and one non-technical (NT7) said that they saw it during study but according to our observation, they didn't mentioned its presence or absence for legitimacy of the website. NT7 mentioned that "I saw it on few banking websites", but didn't mentioned it as a reason for making decision. Participant NT4 started seeing it later in the study and didn't know its exact meaning. One participant, NT10, didn't mention any website fraud/legit on the basis of lock icon; but when asked in the interview, said that "Yes I did, but didn't mentioned".

NT11 didn't check during study and said "Usually I do, especially with banking websites and accounts. It means secure website."

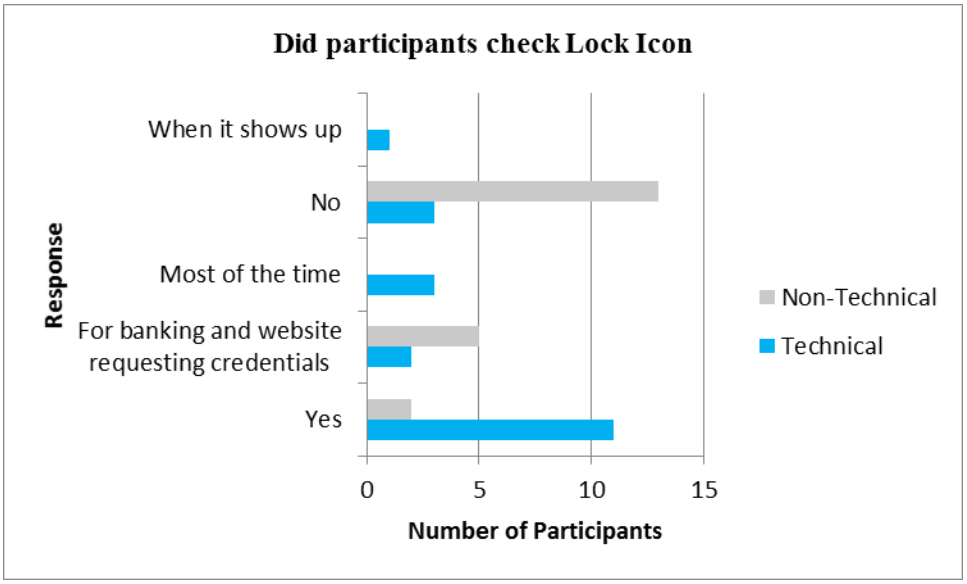

Figure 14 Did participants check Lock Icon

On further probing them for reasons, T14 doesn't look at it in daily life and said "No, I don't have the habit, it needs username and password for accessing a few features of the website". T2 said "my impression is it makes me feel safe. If something is locked it should be safe. Connection private means 70% safe". T18 checked the lock icon for few websites during the study but stated that they check most of the time in daily life. NT14 said "Guess, information you are putting in is secure". NT16 said "I assume safe/ok to use or it has been approved. I don't usually check".

Understanding of domain highlighting:

We asked the meaning of domain highlighting only in the questionnaire. Sixteen technical and eighteen non-technical participants didn't know what it meant. Only one technical participant (T11) gave the correct answer. The other three technical participants (T4, T6 and T12) related it to security and legitimacy of the website. The answer that the two non-technical (NT7 and NT18) gave was incorrect.

Understanding of SSL/TLS certificate:

We asked the meaning of SSL/TLS in the questionnaire. Seven technical and 16 non-technical participants didn't know its meaning. Eleven technical participants related it's meaning to "security and trust towards the website" and the remaining two technical participants (T6 and T16) gave the exact meaning. Out of four non-technical participants, who gave an answer, two (NT2 and NT4) related it's meaning to "security", NT3 said "communication is secure" and NT7 said "its authentication certificate".

When probed in the interview for checking its presence, we had four categories of responses: “yes, with details”, for those who also see the details by clicking on lock icons; “for banking and websites requesting credentials”; “No”; and “checked verified by”, for those who only check that website is verified by which third party or if it is verified or not, as shown in figure 15.

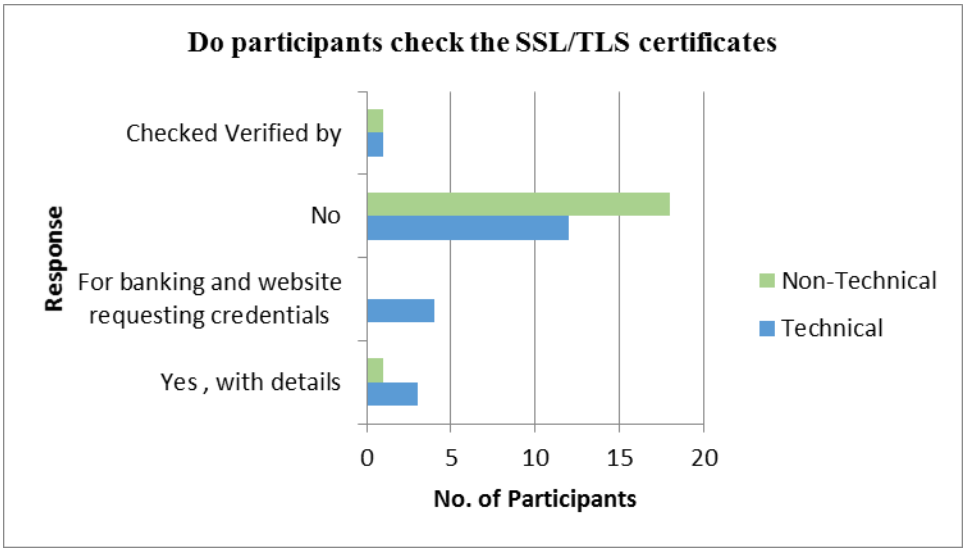

Figure 15 Response of checking SSL/TLS certificates

When technical participants were probed on the reasons for not checking it T1 and T20 said that they “don’t remember how to check that”, T6 said “it doesn’t grab my attention”, T14 and T15 said “I’m not in a habit to check it”, and T17 stated “I forgot to check during study”.

In case of non-technical participants, 17 were not aware about it. NT2 said “I don’t check usually, just check verified by because it am doing a study”, NT3 said “I don’t know how to check”, and NT7 said “no reason for not checking it”.

Looking at URL:

When asked in the questionnaire, 19 technical and 16 non-technical participants said "yes". One technical participant T10 checked only if something looked suspicious. For the other four non-technical participants, two (NT6 and NT20) don't check it and two (NT2 and NT18) stated "sometimes".

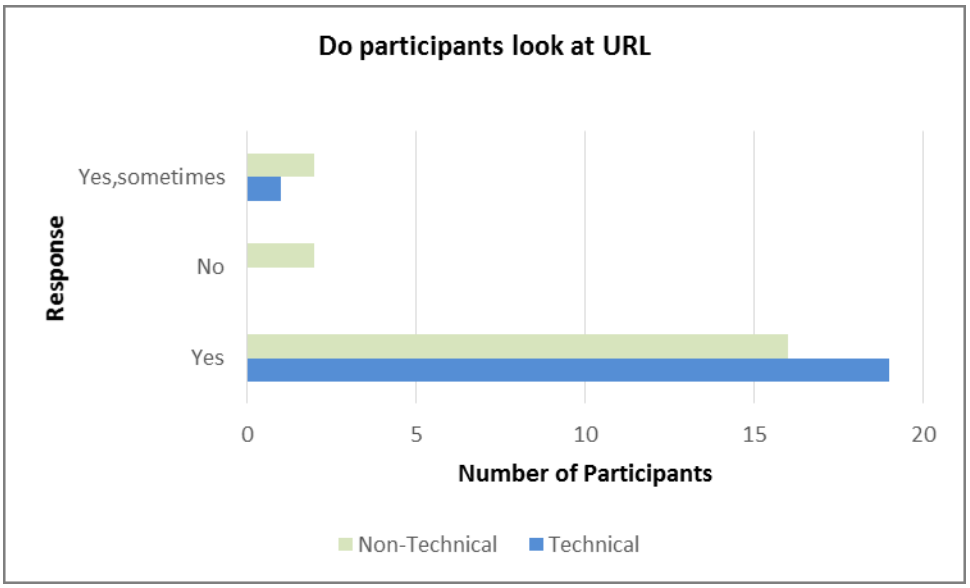

Figure 16 Response of looking at URL

When probed in the interview, six non-technical participant's (NT1, NT6, NT17, NT11, NT12 and NT13) didn't check URL bar during the study with four participants (NT1, NT11, NT12 and NT13) who stated they usually check it in their daily web surfing. When asked for the reasons, we got a variety of responses and figure 17 shows the reasons along with number of participants mentioning those reasons. T1 said "in first website I was not aware about what exactly I was looking for, and also I use that website daily. Then after the first website it occurred that seeing domain name like this may not be legitimate and then you change your behavior, I think I should check that which

website I'm visiting is secure. In general: I do not look often at url, unless it requires banking/personal details."

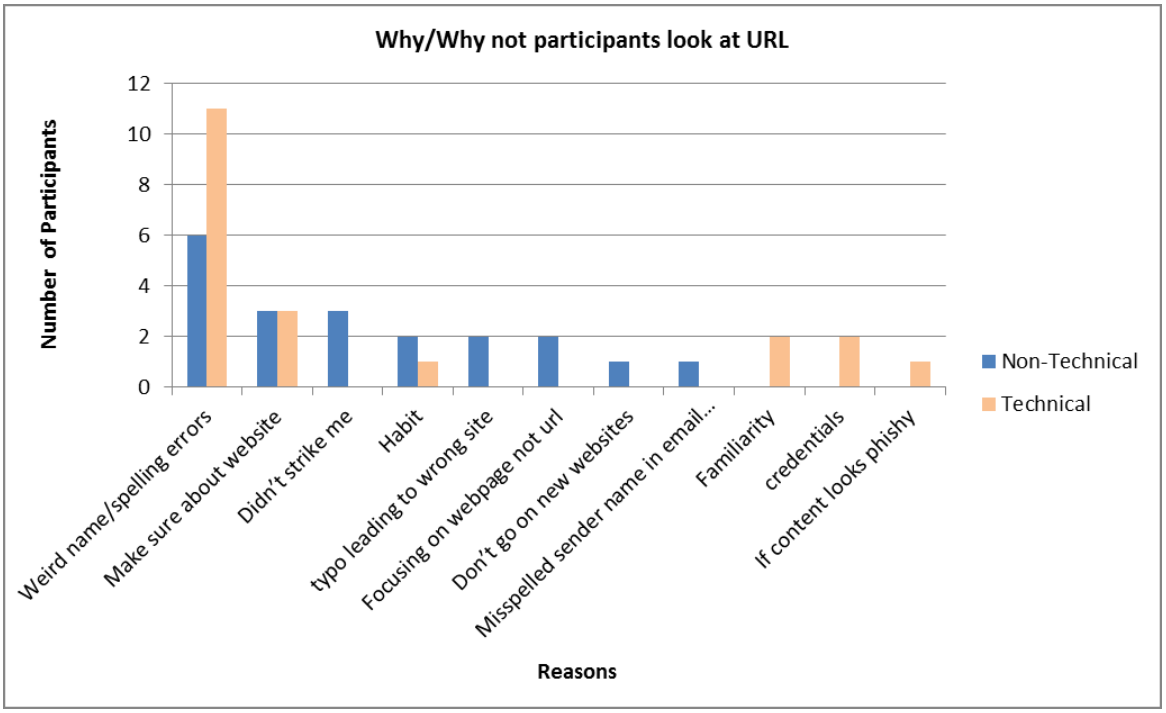

Figure 17 Reasons for looking/not looking at URL

T3 said "In CS we play with the URL's. It's a habit for us what our subpages will be". T6 said "it's the first thing I look at, if spelling is correct then it will be real, if there are any extra symbols or numbers then I check". T8 said "address can be similar to what I am looking for, so I check URL. Weird naming cannot be trusted". T12 said "websites can be mimicked and that can be very confusing, they cannot have same domain name, so I need to check URL".

NT2 said "I have been taken to sites before where I made a typo and I gone to an unintended site, so I check it". NT8 said "sometimes it looks real, but people change domain name. So to make sure I need to see it". NT11 said "I didn't do in the study because I type websites I just didn't think about it". NT14 said "I assumed it was like I

have to ignore URL bar, generally I do, spelling mistakes in address means they are wrong". NT18 said "I looked because I noticed misspelled sender name in email, so it alerted me to see".

Three non-technical participants mentioned that it didn't strike them to see URL bar; out of them NT20 said "It didn't catch my attention because I don't know that by looking at that what's the use of it".

Visiting unfamiliar websites:

Figure 18 describes the response of each participant for visiting unfamiliar websites when asked in questionnaire.

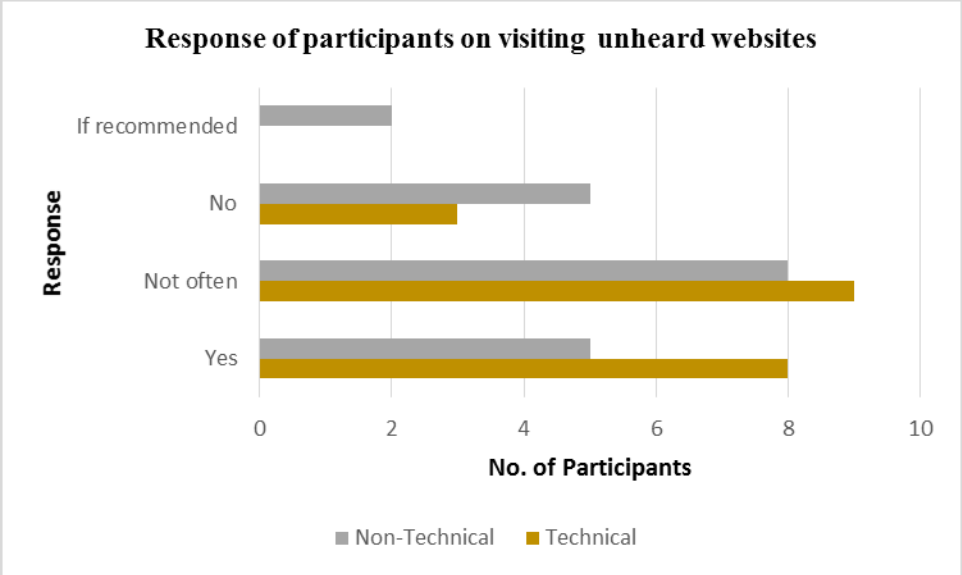

Figure 18 Response on visiting unheard websites

When asked about the reasons for visiting or not visiting those websites in interview, again all the 40 participants gave several reasons, we divided them into categories accordingly and figure 19 shows the number of participants falling into each category. Most common reason given by technical participants included "Google check" and

"reviews". T2 said "you don't know exactly what it is, I will check on google, if it says its good then I check". T9 said "If I searched on google and get them in results then I visit and also I check layout. I go for blogs but not for my credentials". T6 said "after looking at reviews on You Tube and Instagram. Only after trying to search for it first, if it works for others then I try one. Also if it is an advertisement sponsored by Instagram".

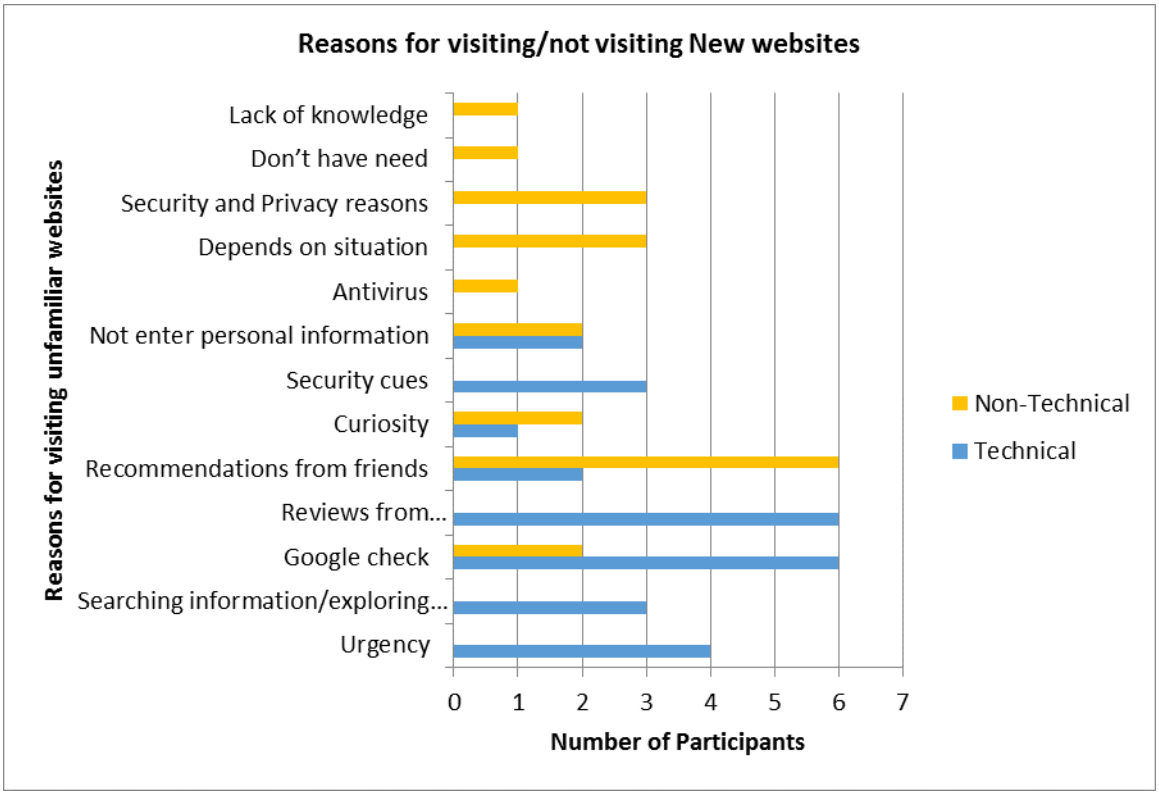

Figure 19 Reasons for visiting/not visiting unheard websites

In case of non-technical participants, the most common reason given was "recommendation from friends", 6/20 participants were willing to visit new websites if they are recommended by their friends, or told by someone to visit them. NT8 said "Sometimes, when I came here then I visited websites which are used here. I listened

from people and if they recommended then. I don't open myself". Two stated that they don't enter personal information unless they verify it as authentic.

Two participants visit out of curiosity. As NT1 said, "If the subject matter looks interesting and depends on my McAfee to figure out if that is dangerous."

Two participants relied on the context of the situation to visit new websites. NT11 said "It depends on a kind of website it is, like for example if I want to watch a movie, I'll try a new website for that or if I am collecting information occasionally".

Four participants do not visit new websites: three out of whom don't visit due to safety reasons and the fourth one because of no need. NT14 and NT18 mentioned they use Google search to visit new websites. One participant reflected out the lack of knowledge, NT20 said "I don't visit much, I keep busy with websites which I frequently visit. I don't know much details to figure out which is good and bad, so I don't take risks."

Entering password in a website reached through an email, given options (Yes often, Yes rarely, No, I don't know):

As seen in figure 20, out of 40 participants, 7 technical and 6 non-technical opted "yes, rarely", 13 technical and 12 non-technical opted "No", two non-technical said "yes often" and "I don't know" respectively. We asked this question only in the questionnaire.

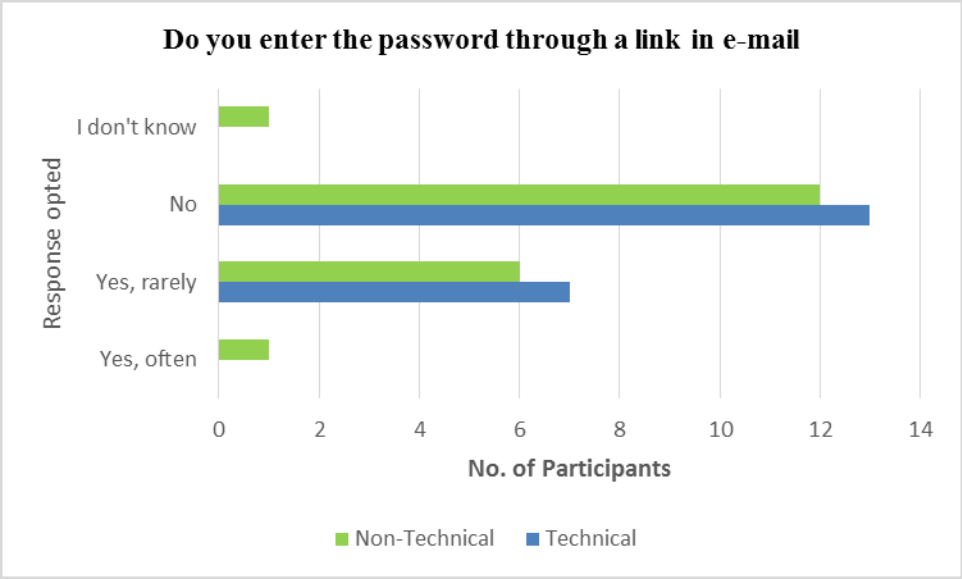

Figure 20 Response for entering password through a link in e-mail

Clicking links in emails received by unknown sender, given options (Yes often, Yes rarely, No, I don't know):

Out of 40 participants, 4 technical and 5 non-technical opted for "yes, rarely", 16 technical and 15 non-technical opted for "No" as shown in figure 21. We did not followed upon this in the interview.

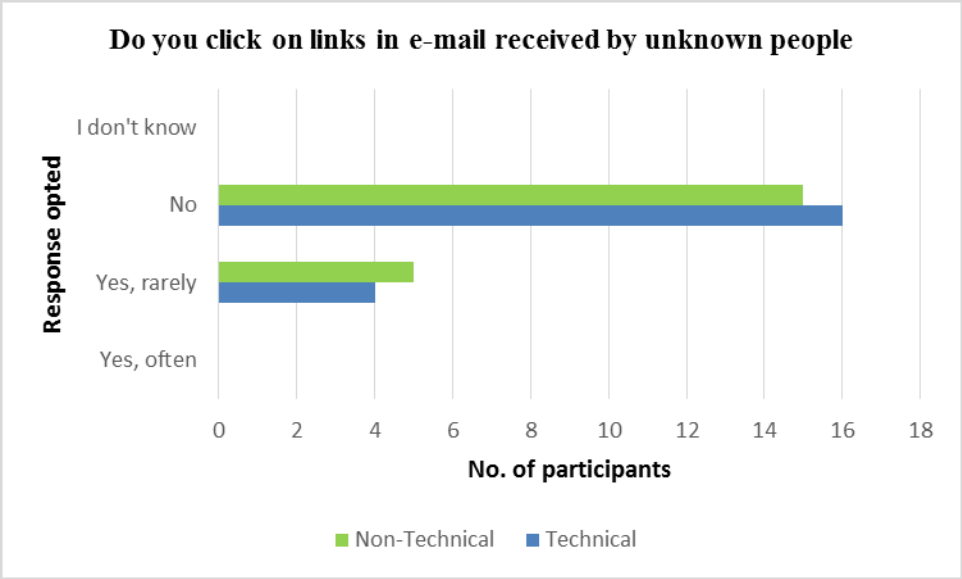

Figure 21 Response for clicking on links in e-mail received by unknown people

Submitting personal information to emails received from personal information, given options (Yes often, Yes rarely, No, I don't know):

All 40 participants opted for option "No" as shown in figure 22. We asked this only in questionnaire.

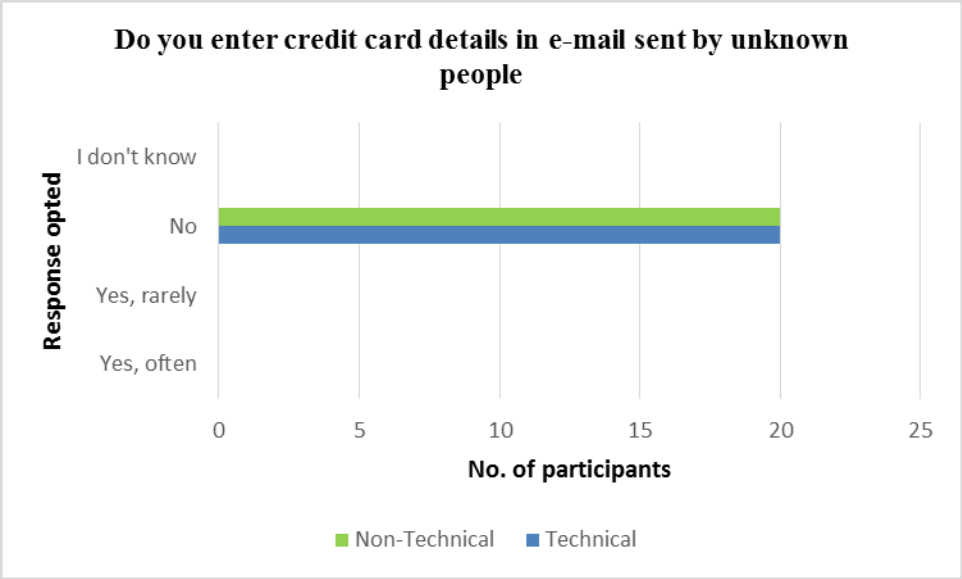

Figure 22 Response for entering credit details in emails received by unknown people

*Suspicion about spelling/grammatical errors in email, **given options** (Yes often, Yes rarely, No, I don't know)*:

All 20 non-technical opted for "yes, often", whereas out of 20 technical participants, 14 opted for "yes, often", 3 said "yes, rarely" and 3 said "No" as shown in figure 23. We asked this only in questionnaire.

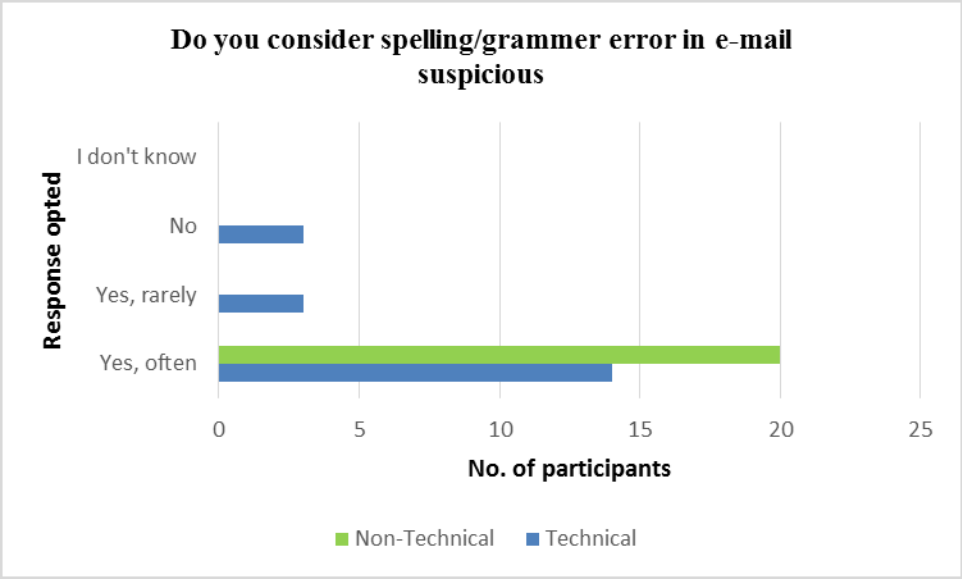

Figure 23 Response on suspiciousness of spelling/grammar error in e-mails

We asked some additional questions in the interview which are described below:

Understanding Phishing:

When participants were questioned about the meaning of phishing, all 40 participants gave varied explanations. We have categorized them and the responses included, “Loss of information“, “misusing information”, “online attack”, “stealing information”, “fake website”, “mass emails”, “wrong things”. The number of participants who responded to a particular option can be seen in Figure 24. NT3 said “I think, wrong URL’s or some emails with link ask you to enter information and steal it”. T8 said “Getting your information and using for different purpose, through emails sending links, faking and fraud”. Eight of the twenty non-technical participants said they don’t know about it. For example, NT9 said “I don’t know”, NT11 said “No Idea”.

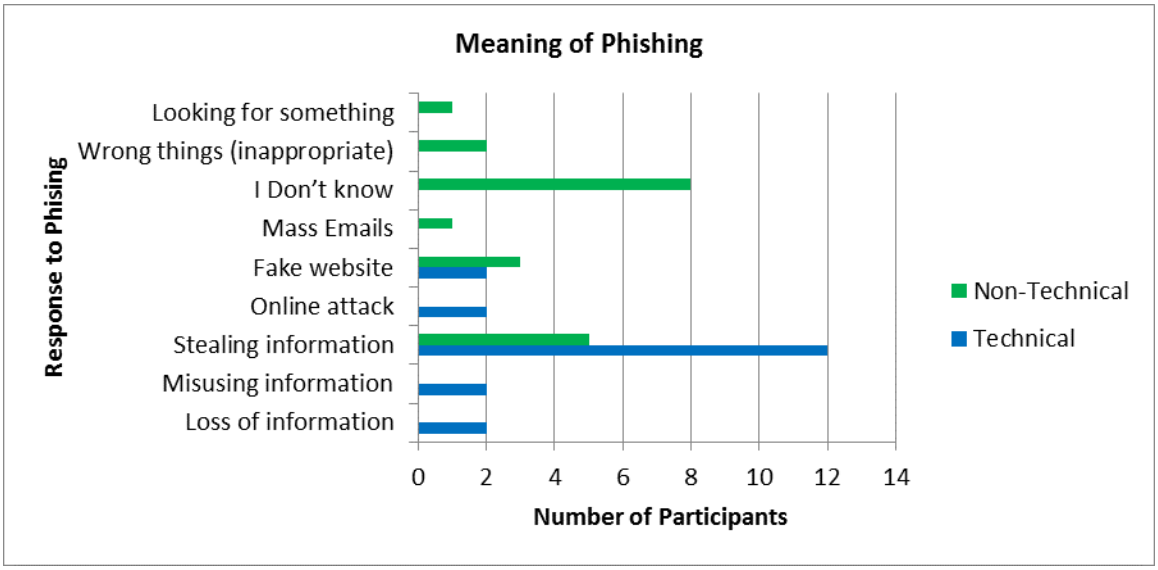

Figure 24 Meaning of Phishing illustrated by participants

Suspiciousness about links in emails:

We also asked this question to the half of the participants who did not see emails during the study (i.e., in general, when they receive emails, what makes them suspicious). The responses we received were either "yes", "sometimes" or "depends on the sender". Figure 25 shows the number of participants for each response.

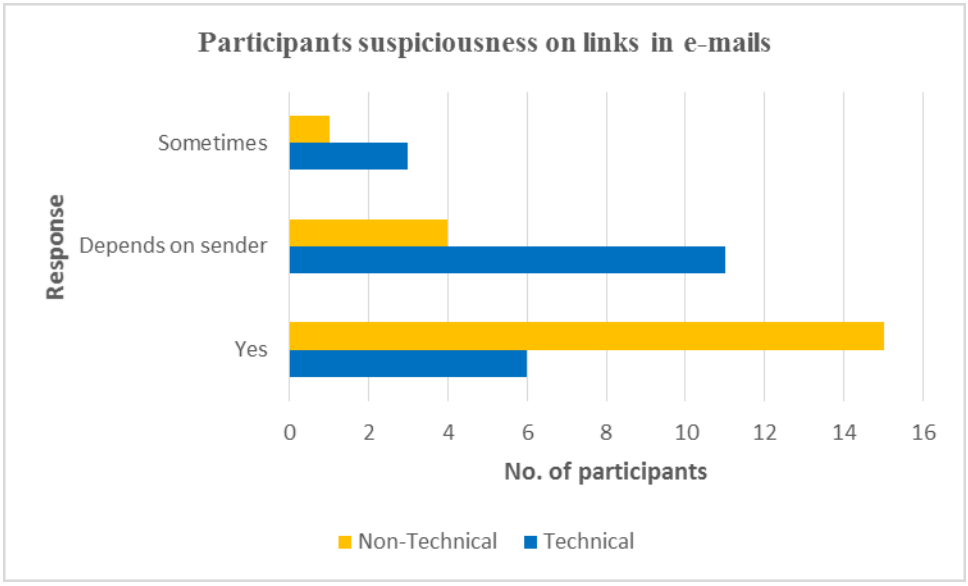

Figure 25 Response to suspiciousness of links in e-mails

When asked what made them suspicious, participants mentioned various characteristics of emails making them suspicious, we categorized them as shown in figure 26. The majority of the participants relied on known/unknown sender for making decisions about the emails. Other features of the email which makes them suspicious included: types of email, organization sending the email, weird link containing letters and numbers, signature of email, content of the email, spelling mistakes, links being different from emails.

Two technical participants (T14 and T15) mentioned that they do not click on the links even if the email seems genuine, rather they go to the websites separately. In case of banking/payment related emails, T15 and T19 said that they contact their banks or companies from which it is coming to confirm about it.

Two other technical participants had different views on clicking link in the emails apart from seeing the above mentioned suspicion features: T16 mentioned that “a few emails will be more professional and even I do click on links, then after proceeding to website only I get to know whether it’s a spam or not”; also T16 stated that “I sometimes opens an email from unknown sender and clicks on link to see what’s there”. T8 said “before I click I hover my mouse and see address, if brand name, I click on it always”.

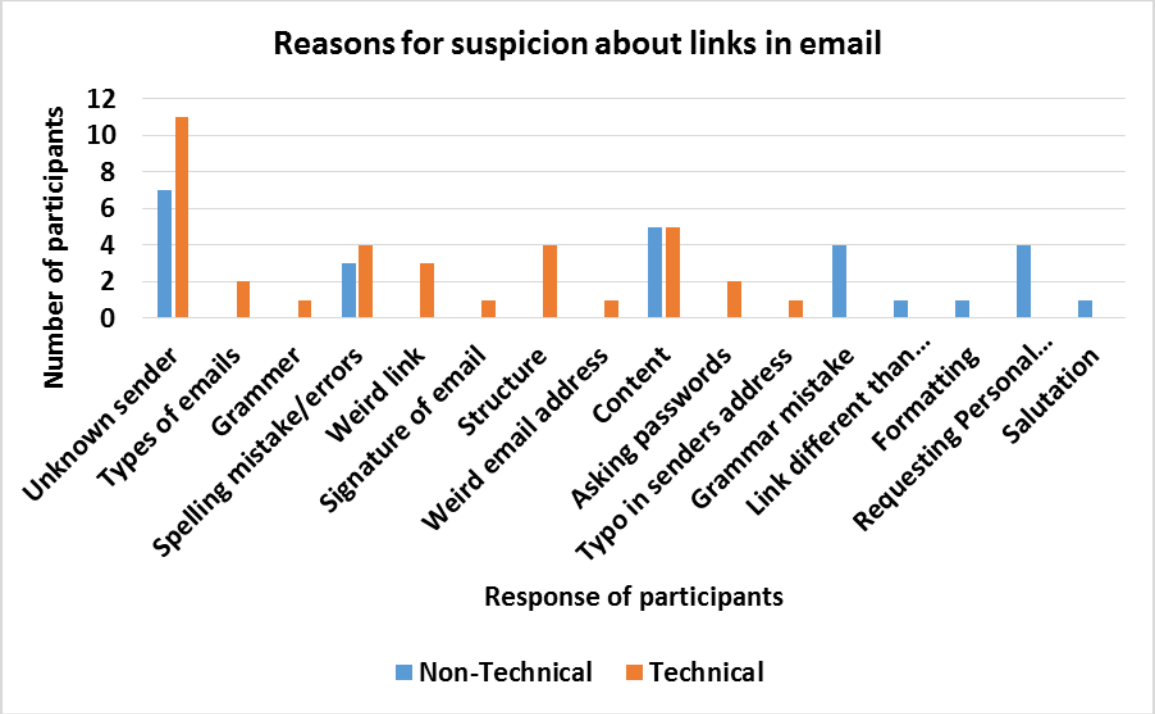

Figure 26 Reasons for suspicion about links in e-mails

One non-technical participant NT2 said "I'm suspicious about all unnecessary emails, anytime I receive emails from company I haven't signed up for like news, offers anything unknown then I'm suspicious. If there is something I might be interested in, I take a look on link, if link appears suspicious, if it says buying chocolate but link says findhotguys.bz or something I will just delete that."

NT4 said they were often suspicious but still click on links to see what's there. NT20 stated they don't trust emails usually and said "If it is urgent, I call customer service or use mobile OTP".

Website's important features against phishing attacks:

Figure 27 shows top features of website considered important by both technical and non-technical participants.

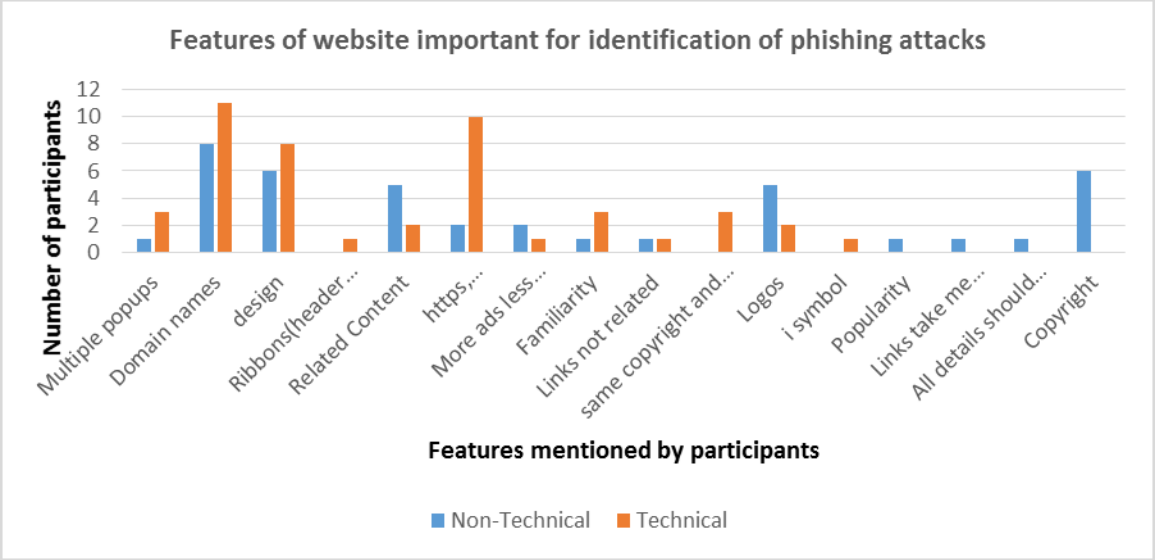

Figure 27 Website features mentioned important by participants for identifying phishing attacks

*Factors influencing decision on websites (1: no influence to 5: strongly influence), (factors given: Content, URL, Look and Feel and other factors)**:***

Figure 28 shows the influence of the above given factors on both the technical and non-technical participants in making decision about the website.

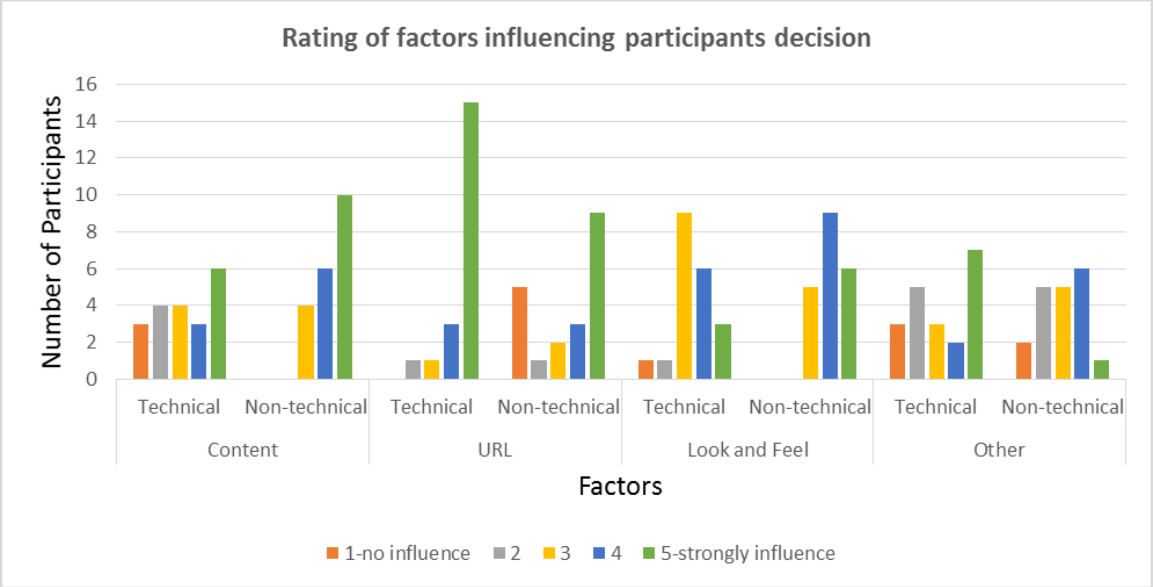

Figure 28 Rating of factors influencing participant's decision

Effectiveness of security cues with improvements:

When we asked participants about the effectiveness of browser security cues, 25/40 mentioned either they are not effective enough or only to some extent and sometimes as shown in figure 29. T2 said "Effective sometimes, if Google says no, I won't go to check security cues". T9 said "If person is knowledgeable, then only they are effective". 6/20 technical and 2/20 non-technical said yes they are effective. Whereas, only 3 participants (2 non-technical and 1 technical) said that they are not effective. Also four non-technical participants (NT1, NT11, NT18 and NT20) said that they don't pay attention to them so they cannot say about its effectiveness. NT11 said "I don't usually look on those, I have to download bitdefender, ad blocker, which gives me a message". NT20 said "No, I don't know much about it."

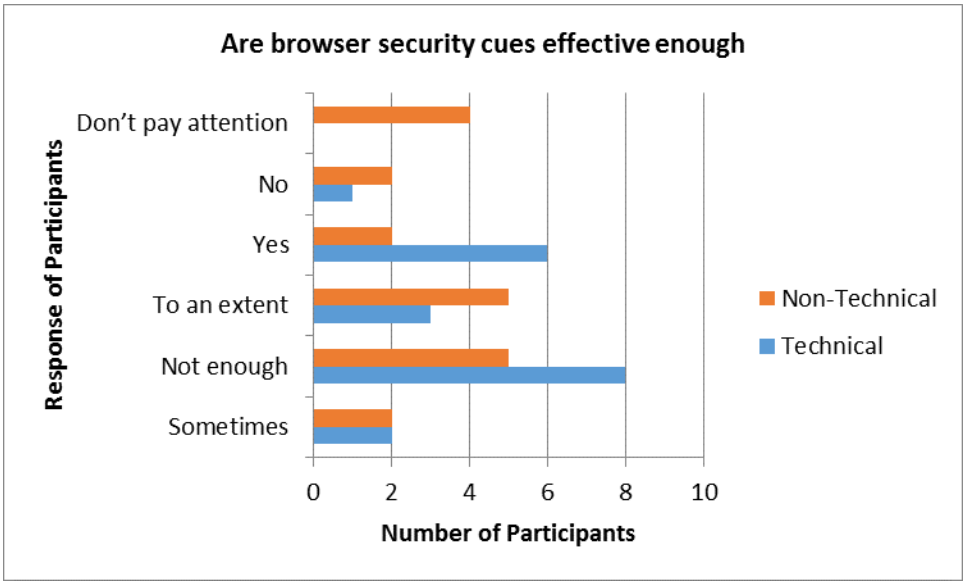

Figure 29 Effectiveness of browser security cues as per participants

We also asked participants for suggestions and improvements in the design of security cues. The suggestions are as follows:

1. **Popup showing security information (10 Technical and 4 Non-technical):**

   14 participants mentioned that pop-ups displaying all the information about the website security will be great help. T5 said "As soon as you visit a website, there should be a pop up showing that website is not legitimate. It should be like an anti-virus." T7 said "if there will be something which shows google ranking of the website, no. of visits to that website, or something new (something in browser which directly makes websites as spam)".

2. **Symbolic Representation of security on website (4 Technical and 3 Non-technical)**

   In this category participants mentioned to improve the overall representation of security cues and to include *hover help* as well. For example, T10 said "To simplify these cues, put your cursor over them and get information about what they are". T14 said "there should be a balloon in the corner saying this is https or locked website".

3. **User Education (2 Technical and 4 non-technical)**

   Six of the 40 participants felt it was necessary to train and make users aware about important security cues, as most of them don't pay attention to them due to lack of knowledge. For example, NT12 said "more awareness should be there, letting everybody know about the checklist like you have to check URL, etc." NT15 said "I don't know if Https is supposed to be there only for banking websites, I wish I had that knowledge".

4. **Representation by colour (1 Technical and 1 Non-technical)**

   In this category, participants mentioned using different colour coding scheme to represent security symbols. For example, NT19 said "have URL of different colors on certified websites". T16 said "if browser finds something suspicious it can have yellow thing along with green and red".

5. **Third party verification/validation (1 Technical and 1 Non-technical):**

   There should be a way to verify and validate security for a website using third party rather than user itself checking it. For example, T1 said "as a user has a connection to 3rd party who are protecting your information, you should not check on your own, 3rd party should do the verification. Certificate is hidden which doesn't remind to check it and its time consuming". T3 said that there should be some kind of CAPTCHA system that verifies the website legitimacy.

6. **Making certificate details visible (1 technical and 1 non-technical)**

   Two participants mentioned that certificate information should be visible. T18 said "certificate should be made visible because I didn't know how to use it".

7. **Favicon in URL (1 non-technical):**

   One participant NT10 said to include favicon in URL on google chrome as it is represented in Internet explorer. NT10 stated "I go to websites through google chrome. Logo of website on URL should be there as in Internet Explorer".

8. **Synchronisation of browser history with email (1 technical)**

   One participant T20 suggested that there should be a way for browsers to synchronize their browser history with users email account. For example, if user accessed 20 websites, browser should notify user through an email about the

legitimacy of the accessed websites by sending an email. T20 stated "when someone use websites, based on browser history an email should be sent about security".

Things to keep in mind in order to protect from phishing websites and e-mails:

In this question, we asked participants about what security guidelines they follow to protect themselves from phishing attacks. Table 18 and 19 lists the categorized responses from all the participants.

Table 18 Online practices followed by participants for protection against phishing websites

Responses for websites	**Technical Participants (n=20)**	**Non-technical participants (n=20)**
Using antivirus	6 (30%)	4 (20%)
Not visiting unfamiliar financial websites/asking credentials	2 (10%)	7 (35%)
Websites with pop-ups/ads	1 (5%)	4 (20%)
Prefer phone/personal banking	1 (5%)	2 (10%)
Relying on URL and website content	3 (15%)	0
Not saving credit card details online	1 (5%)	1 (5%)
Don't trust anything	2 (10%)	0
Deleting cookies and history every day	1 (5%)	1 (5%)
Using incognito	1 (5%)	0
Checking before clicking	1 (5%)	0
Logging in via fb	1 (5%)	0
Checking websites by entering fake passwords	1 (5%)	0
Using complicated passwords	0	1 (5%)

For maintaining safety with emails, 13/40 participants said that they don't trust emails from unknown sender, also 2 of them either unsubscribe them or mark it as spam. Six of them preferred going to websites separately rather than clicking on link provided in the email and leading to the website. Five participants were very cautious about emails related to financial information and before proceeding with it always want to contact bank to be sure about it.

Table 19 Online practices followed by participants for protection against phishing e-mails

Responses for emails	**Technical Participants (n=20)**	**Non-technical participants (n=20)**
Not trusting emails from unknown sender-mark it spam	8 (40%)	5 (25%)
Not going to website through email	3 (15%)	3 (15%)
Emails with financial information-confirm with bank	2 (10%)	3 (15%)
Don't trust anything	2 (10%)	0
Content	4 (20%)	0

Question: *Are you aware of any phone phishing scams? How do you judge it as legitimate or not?*

28/40 participants (11 NT and 17 T) were aware about these phone phishing scams as can be seen in figure 30.

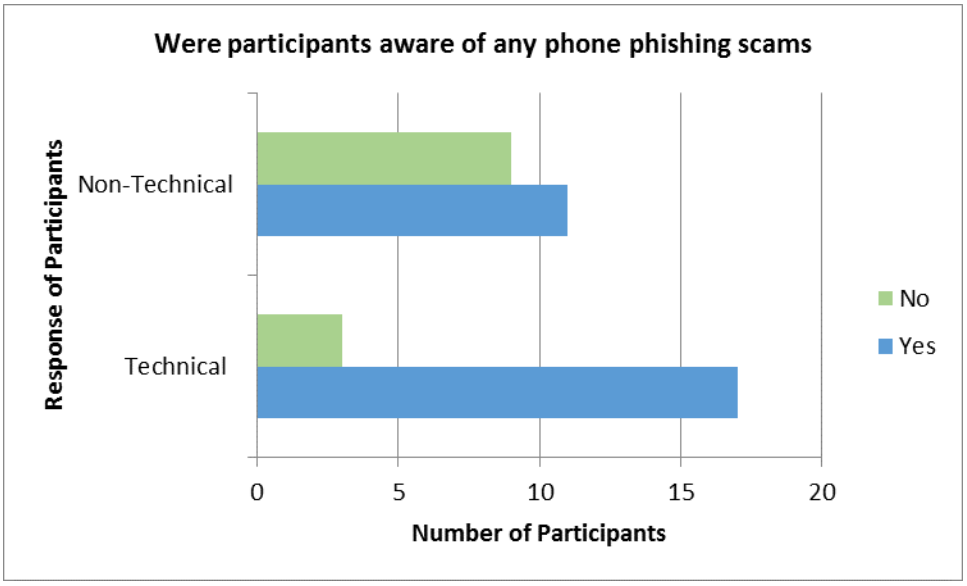

Figure 30 Were participants aware of phone phishing scams?

They either experienced it themselves or they knew it through their friends. When asked how they judge them as true or not, we received variety of responses and after categorizing them we came up with main ones as shown in figure 31. As an example T7 stated “I have “true caller” app, it shows how many people mark it as spam, so it gives me a hint”, T12 said “I usually check if person is in hurry, is there any threat from conversation or is there any money information involved”, T16 said “I ask their identity, go to that website and check if they are calling or send them an email and call them if they are real”. NT10 stated “I saw a video, which shows if you click on link received in text messages. It gives them every information about your phone. I don’t think phone calls can do severe harm”.

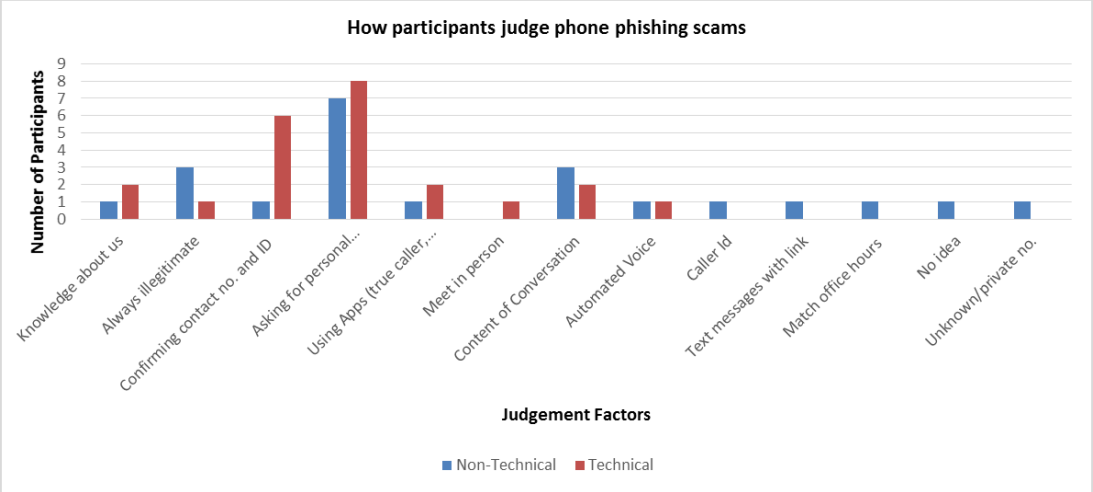

Figure 31 Judgement factors used by participants for phone phishing scams

CHAPTER 5 DISCUSSION AND IMPLICATIONS FOR DESIGN

In this chapter, we discuss important points derived from our findings (Section 5.1) and will provide recommendations for designers and developers of website security cues (Section 5.2).

5.1 DISCUSSION

5.1.1 Lack of Security Awareness

The results in Section 4.1 and 4.4 suggest that a lack of security awareness makes users more prone to phishing attacks. As can be seen from the interview results in section 4.6, 40% of non-technical participants didn't know what phishing means, 65% of them didn't look at the lock icon, out of which 46% didn't know about what it indicates and 19% answered it incorrectly. Also, 85% of the non-technical participants didn't have any knowledge about SSL certificates. This indicates that there is still need for increasing awareness about security indicators, which could help users to locate their presence or absence in identifying fraudulent/legitimate websites.

5.1.2 Differences between Technical and Non-Technical Participants

From Section 4.1, we found that there were significant differences between the technical and non-technical participants on the basis of accurateness and dangerous mistakes committed while judging the websites. Technical participants performed better than non-technical participants in identifying phishing websites. Also, Section 4.3 revealed different strategies followed by the participants, out of which a few were specifically used by technical group including, hovering over separate links and comparing address

generated from them to the domain name of the website, checking Meta data in inspect element, checking source of the page and comparing copyright spelling with the domain name. Whereas, one of the strategies (i.e., recommendations from friends) was used by only one non-technical participant. This suggest that, if non-technical users also start hovering over links to compare URL's they would be able to recognize more phishing websites.

5.1.3 Familiarity

As compared to the findings of Kelley and Bertenthal [13] in terms of familiarity of the website, we also found that participants tend to ignore security cues if they are very familiar to the websites. For example, when we showed my.da1.ca to our participants, two of the participant's (NT9 and T18) were ready to proceed with the login even after noticing misspelled domain names. NT9 stated "*should I look at my.da1.ca or is it ok? It looks real*" and T18 was ready to give credentials believing that it was from university and said "*it might be a mistake by them*" when noticed "1" in domain name instead of "l". This again suggests a great need to improve security indicators and increase user knowledge.

5.1.4 Differences based on Gender

Alsharnouby et al. [7] and Dhamija et al. [6], didn't find any significant difference based on gender in overall score. However, when we compared the scores for identifying correct phishing and legitimate websites, we were able to find significant differences between males and females as can be seen in Section 4.1.2. Non-technical female participant's committed more dangerous mistakes by assessing phishing websites as

legitimate as compared to male participants. According to the findings of Sheng et al. [39], as compared to male participants, females who lack security awareness in terms of technical knowledge were more prone to phishing attacks and Jagatic et al. [24] also found that the percentage of females (77%) who fell for the phishing e-mail attack was more than that of males (i.e. 65%). This indicates that females are likely to fall for phishing attacks as compared to male participants.

5.1.5 Using own laptop

During our study, the majority of participants used their own laptop to perform the experiment as compared to previous studies (i.e., [6 and 7]) that used lab computers. We were able to find that if users were already logged in to the websites or their login information was saved by their browsers, they answered the website as legitimate (example, Facebook, amazon, LinkedIn, etc.). Two participants T20 and NT6 mentioned the website as fake if they not find themselves logged in to the websites in which they were supposed to be. This strategy i.e. "I am logged in or not", could not be observed when lab computers are used for these experiments. Also, if we compare paying attention to security cues or participant's performance, we got similar results as compared to [6 and 7]. We got an average success rate of 74.7 % (both technical and non-technical participants) in comparison with 64% for Alsharnouby et al. [7] and 58 % for Dhamija et al. [6], whose participants appear to be primarily non-technical in nature. The slight improved results is because we included both technical and non-technical participants. Our technical participants performed significantly better than the non-technical participants. This may indicate that participant's don't pay extra attention to security cues when using their own laptops for this type of experiment.

5.1.6 Using Browsers of Participants choice

As compared to previous studies [6, 7 and 27], we asked participants to perform the experiment on the browser of their own choice, which they use on daily basis. There were a couple of participants (T16 and NT9) who stated that they used multiple browsers on daily basis. So, even having chosen a browser themselves, they were still confused with some of the security cues. Participant T16 used Safari and answered one legitimate website as phishing because of grey lock instead of green. Similarly, NT9 mentioned on one website that favicon should be in the URL, not as it is currently appearing. NT9 confused this feature with its location on Internet Explorer. This indicates that the choice of browser can have an impact on the decisions made by the participants.

5.1.7 Changes in chrome SSL indicators

Starting from January 2017, there has been some changes in the display of security indicators in Chrome. The word "secure" now appears in addition to green lock icon, as shown in figure 32. And the words "not secure" are used on sites that collect user credentials without Https.

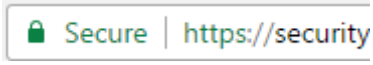

Figure 32 New feature in Chrome [37]

Currently, all sites without Https do not have "not secure" on them. For example, indicators on a website containing fields requesting credit card details is shown in figure 33 [37]. It doesn't have any warning or not secure symbol and can fool those users who don't pay attention to security cues.

Figure 33 Presence of indicators without Https [37]

Many people still get panicked when their screen displays a pop-up saying that a virus has been detected and they need to clean their system by calling the number provided. So, research is ongoing in this area to improve these security indicators and thus it is a "hot research topic".

5.2 RECOMMENDATIONS

In this section, we suggest some recommendations which might help the designers of web browsers security cues to improve them and thus help users in identifying phishing websites more exhaustively. On the basis of feedback received from the participants, our results and discussion, we provide following recommendations:

5.2.1 Box of Security information on Website

Many (35%) of our participants, as mentioned in Section 4.4, showed an interest in including a box as a pop-up or a balloon containing website information such as whether website has a valid certificate or not, secure to use or not, its google ranking, etc. The presence of a separate box as a pop-up or balloon before a website is about to open will help direct users attention to it as is also mentioned by Egelman et al. [10] in their recommendations about importance of active warnings that interrupt user tasks. And also from our findings, when participants were asked about the reasons for not looking at security indicators, some participants said that it didn't caught their attention to look at them.

5.2.2 Hover Help

As some participants are not actually aware about the meanings and purpose of security cues as found in section 4.4. We consider that it will be helpful to include hover help on the security cues. For example, when user hover over https or lock icon or URL, the purpose and meaning of that indicator should appear to help user understand and make use of them properly.

5.2.3 Using Color Scheme for URL

Some participants assessed phishing website as legitimate as a result of landing on the original webpage, which they were directed to from phishing website by clicking on any of the links. This was similar to the findings of Alsharnouby et al. [7] and it revealed that those participants didn't noticed changing URL's while transitioning from a fraudulent page to the original one. Alsharnouby et al. [7] suggested to include an indicator informing users about this transition, but they mentioned that application of this indicator might create difficulties in navigation.

So, we suggest to include color coding scheme for URL's, also suggested by one of our participant in section 4.6. According to this scheme, the domain names of the verified and secure website should be of different color, for example, blue as compared to the domain names of the websites which are not verified but are still legitimate, for example, yellow and finally for rest of the websites which are unsafe, domain name can be of red color. This will also help when there is a transition to or from one domain to a different one as color of URL will change which will help to call attention to it. We believe that this scheme will help users to grab their attention towards the URL and adding hover over option to this will help them understand their meaning and purpose.

In addition to this, we also found that none of our participants knew about "domain highlighting" when asked in post observation questionnaire. No participant mentioned it during the whole study and it was confirmed during the debriefing session that it was new for them. So, this color coding scheme might help making domain highlighting more prominent.

5.2.4 Security Cues per Browser

As mentioned in Section 5.1.6, two participants were confused with the security indicators while using one of the browsers they chose. So we suggest to include basic indicators in a similar manner on all the browsers and if not possible, again to mention the meaning of indicators through the above mentioned recommendations. Whalen and Inkpen [3] also recommended to standardize the layout of security indicators from their findings.

5.2.5 User Education

As mentioned in Section 4.6, 15% of our participants mentioned the need of training and raising awareness about security cues. From the results of section 4.6, lack of knowledge is the reason for not paying attention to security indicators by some of the participants and through our analysis. Hence, we recommend to improve user's knowledge on these concepts by including security related courses in the syllabus of non-computer science fields.

CHAPTER 6 LIMITATIONS AND CHALLENGES

In this chapter, we discuss the limitations of our study and challenges faced during the research.

6.1 USE OF EYE TRACKER

We wanted to use eye tracker for the observation session for tracking participant's gazes to make sure that they looked at security indicators. For this we tested eye tracker on our own laptop but it took long time to install and didn't collect data properly. So, it would have been very difficult for us to install the eye tracker software on each participant's laptop due to technical limitations and we decided to make use of video cameras and talk aloud protocol instead of an eye tracker software.

6.2 HOSTING WEBSITES

For showing the websites to participants, we created phishing websites and hosted them through a local lab server. For one of those websites we used IP address of our server as the domain name to implement one of the phishing techniques, similar to one of the websites shown by Alsharnouby et al. [7]. As per our understanding, it was all internal but our server got flagged for hosting a phishing website. Finally, we applied Firewall rules to our server so that the websites can only be accessed through our lab Wi-Fi. Also, we tried to use the domain names of the phishing websites as similar as possible to the original, so we bought one domain name similar to a banking website. We bought it because it was available and didn't check the e-mail account for a while which was associated with it. We received notices from a third party for using that domain name, which we then deleted from our account and transferred to them just after seeing that e-

mail. So, we would recommend to delete the domain names as soon as after the study is completed.

6.3 Participants

Out of 40 of our participants, 35 were university students with nine students working part – time. We were unable to recruit more professionals to take part in the study. The technical participants were not security experts/professionals but were from computer science background and we used a screening questionnaire to recruit them accordingly. Also, our non-technical participants consisted of highly educated population otherwise, we may have been able to find more differences.

6.4 Not Using Participants e-mail accounts

During observation session, we showed emails as web pages to participants and did not use their own accounts for sending e-mails. This might have impacted their behavior but if we would have sent e-mails to their accounts during the study, that would again be biased because participants would know that those e-mails were sent for the purpose of study and thus lack realism. We also didn't want the e-mail account to flag anything to make it spam. If the mail server is doing a good job it should have noticed and would not let later participants receive that e-mail. This is also a reason for not using participant's actual accounts in this type of study.

6.5 No credentials

In starting of the study we instructed participants to not enter their credentials to login to the websites. This was because of ethical and privacy concerns, as we would have otherwise collected a great deal of participant's sensitive data.

6.6 Lab Environment

As the study was conducted in a lab environment, this artificial environment could have let participant's feel safer than they might have during their normal activities.

6.7 Study Context

Similarly the context of the study, where they were instructed to assess the trustworthiness of the websites is not representative of real world situations where security is a secondary task our results therefore are a best case scenario and we would expect accuracy to decrease when security is not the primary task.

CHAPTER 7 CONCLUSION AND FUTURE WORK

In this chapter, we give the conclusion (Section 7.1) and identify the possible future work (Section 7.2) that can be further carried out based on this research.

7.1 CONCLUSION

Findings from our study reveals that the average success rate for correctly identifying phishing websites is 67.3% for both technical and non-technical participants, even when participants use their own laptops instead of lab computers to participate in the experiment. Less than 50% of our non-technical participants looked at SSL indicators to make decision about the websites. Thus we find that even on their own laptops participants don't pay extra attention to browser security cues. But more than 50% of the participants made use of the strategy "I am logged in or not" to identify the legitimacy of the websites, which can only be explored if participants use their own laptop for the study.

There were significant differences in identifying phishing websites correctly based on technical expertise and Gender. The technical participants with average success rate of 79.2% performed better than non-technical with average rate of 55.4% and Male participants performed better with average success rate of 74.6% than the Females with 57.9%. There were a few strategies that were solely used either by some of the technical participants (i.e. hovering over links to compare url address generated from them to the domain name of the website, checking meta data in inspect element and comparing copyright spelling with domain name) and by non-technical only (recommendations from friends). Also two of our participants were confused with presentation of security cues

because of using multiple browsers on daily basis and thus we found that the browser used during study also has an impact on the judgement of websites.

Of the 20 participants who visited websites through emails, ten of our participants who were suspicious about e-mails judged the websites linked through those e-mails also as suspicious. In the end, we provide some recommendations which might help in the improvement of browser security cues and increase user awareness in identifying websites that are not secure.

7.2 Future Work

In future we would like to recruit participants who are professionals rather than university students. We might test users on tool bars implemented with recommended changes and further propose evaluated and effective indicators. In this study we didn't allow participants to enter their credentials and they didn't use their own e-mail accounts. In future studies we would like them to use their own credentials and e-mail accounts to see the e-mails and judge them. We would also like to conduct it as a field study rather than a lab one and also include location based context factor. Finally this study can be done again to test the user's attention on the changed indicators (i.e. word "secure" in addition to green lock) being employed in Chrome 56 from January 2017.

REFERENCES

[1] What is Phishing? | Phishing.org. Phishing.org. Retrieved 19 May 2016, from http://www.phishing.org/what-is-phishing/

[2] Phishing Attack Prevention: How to Identify & Avoid Phishing Scams. Digital Guardian. Retrieved 19 May 2016, from https://digitalguardian.com/blog/phishing-attack-prevention-how-identify-avoid-phishing-scams

[3] Whalen, T., & Inkpen, K. M. (2005, May). Gathering evidence: use of visual security cues in web browsers. In Proceedings of Graphics Interface 2005(pp. 137-144). Canadian Human-Computer Communications Society.

[4] Herzberg, A. (2009). Why Johnny can't surf (safely)? Attacks and defenses for web users. Computers & security, 28(1), 63-71.

[5] Felt, A. P., Reeder, R. W., Ainslie, A., Harris, H., Walker, M., Thompson, C., ... & Consolvo, S. (2016). Rethinking Connection Security Indicators. In Twelfth Symposium on Usable Privacy and Security (SOUPS 2016).

[6] Dhamija, R., Tygar, J. D., & Hearst, M. (2006, April). Why phishing works. InProceedings of the SIGCHI conference on Human Factors in computing systems (pp. 581-590). ACM.

[7] Alsharnouby, M., Alaca, F., & Chiasson, S. (2015). Why phishing still works: user strategies for combating phishing attacks. International Journal of Human-Computer Studies, 82, 69-82.

[8] Downs, J. S., Holbrook, M. B., & Cranor, L. F. (2006, July). Decision strategies and susceptibility to phishing. In Proceedings of the second symposium on Usable privacy and security (pp. 79-90). ACM.

[9] Downs, J. S., Holbrook, M., & Cranor, L. F. (2007, October). Behavioral response to phishing risk. In Proceedings of the anti-phishing working groups 2nd annual eCrime researchers summit (pp. 37-44). ACM.

[10] Egelman, S. (2009). Trust me: Design patterns for constructing trustworthy trust indicators. ProQuest.

[11] Thurlby, C., Langensiepen, C., Haggerty, J., & Ranson, R. (2015). Understanding User Knowledge of Computer Security and Risk: A comparative Study. In Ninth International Symposium on Human Aspects of Information Security & Assurance (HAISA 2015)

[12] Fagan, M., & Khan, M. M. H. (2016). Why Do They Do What They Do?: A Study of What Motivates Users to (Not) Follow Computer Security Advice. In Twelfth Symposium on Usable Privacy and Security (SOUPS 2016).

[13] Kelley, T., & Bertenthal, B. I. (2015). Tracking Risky Behavior On The Web: Distinguishing Between What Users Say'And Do'. In Ninth International Symposium on Human Aspects of Information Security & Assurance (HAISA 2015), S. Furnell and N. Clarke, Eds., no. HAISA. Lesvos, Greece: CSCAN Open Access Repository (pp. 204-214).

[14] Kelley, T., & Bertenthal, B. I. (2016). Real-World Decision Making: Logging Into Secure vs. Insecure Websites.

[15] Jakobsson, M., Tsow, A., Shah, A., Blevis, E., & Lim, Y. K. (2007, February). What instills trust? a qualitative study of phishing. In International Conference on Financial Cryptography and Data Security (pp. 356-361). Springer Berlin Heidelberg.

[16] Parsons, K., Butavicius, M., Pattinson, M., Calic, D., Mccormac, A., & Jerram, C. (2016). Do Users Focus on the Correct Cues to Differentiate Between Phishing and Genuine Emails?. arXiv preprint arXiv:1605.04717.

[17] Ion, I., Reeder, R., & Consolvo, S. (2015). "... no one can hack my mind": Comparing Expert and Non-Expert Security Practices. In Eleventh Symposium On Usable Privacy and Security (SOUPS 2015) (pp. 327-346).

[18] Akhawe, D., & Felt, A. P. (2013). Alice in warningland: A large-scale field study of browser security warning effectiveness. In Presented as part of the 22nd US

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

[20] Shi, P., Xu, H., & Zhang, X. L. (2011, February). Informing security indicator design in web browsers. In Proceedings of the 2011 iConference (pp. 569-575). ACM.

[21]Kurtenbach, G. P. 1993. The design and evaluation of marking menus. University of Toronto, 1993.

[22] Neupane, A., Rahman, M. L., Saxena, N., & Hirshfield, L. (2015, October). A Multi-Modal Neuro-Physiological Study of Phishing Detection and Malware Warnings. In Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security (pp. 479-491). ACM.

[23] Wu, M., Miller, R. C., & Garfinkel, S. L. (2006, April). Do security toolbars actually prevent phishing attacks?. In Proceedings of the SIGCHI conference on Human Factors in computing systems (pp. 601-610). ACM.

[24] Jagatic, T. N., Johnson, N. A., Jakobsson, M., & Menczer, F. (2007). Social phishing. Communications of the ACM, 50(10), 94-100.

[25] Wright, R. T., & Marett, K. (2010). The influence of experiential and dispositional factors in phishing: An empirical investigation of the deceived. Journal of Management Information Systems, 27(1), 273-303.

[26] Sobey, J., Biddle, R., van Oorschot, P. C., & Patrick, A. S. (2008, October). Exploring user reactions to new browser cues for extended validation certificates. In European Symposium on Research in Computer Security (pp. 411-427). Springer Berlin Heidelberg.

[27] Kelley, T., Camp, L. J., Lien, S., & Stebila, D. (2012, July). Self-identified experts lost on the interwebs: The importance of treating all results as learning experiences. In Proceedings of the 2012 Workshop on Learning from Authoritative Security Experiment Results (pp. 47-54). ACM.

[28] Lien, S., Kelley, T., & Camp, L. J. Expert and Non-expert Viewing Patterns When Seeking Security Cues.

[29] Arianezhad, M., Camp, L. J., Kelley, T., & Stebila, D. (2013, February). Comparative eye tracking of experts and novices in web single sign-on. In Proceedings of the third ACM conference on Data and application security and privacy (pp. 105-116). ACM.

[30] Furnell, S. M., Jusoh, A., & Katsabas, D. (2006). The challenges of understanding and using security: A survey of end-users. Computers & Security, 25(1), 27-35.

[31] Wang, J., Herath, T., Chen, R., Vishwanath, A., & Rao, H. R. (2012). Research article phishing susceptibility: An investigation into the processing of a targeted spear phishing email. IEEE transactions on professional communication, 55(4), 345-362.

[32] Friedman, B., Hurley, D., Howe, D. C., Felten, E., & Nissenbaum, H. (2002, April). Users' conceptions of web security: a comparative study. In CHI'02 extended abstracts on Human factors in computing systems (pp. 746-747). ACM.

[33] Friedman, B., Howe, D. C., & Felten, E. (2002, January). Informed consent in the Mozilla browser: implementing Value-Sensitive Design. In System Sciences, 2002. HICSS. Proceedings of the 35th Annual Hawaii International Conference on (pp. 10-pp). IEEE.

[34] Associates, C., Types of Phishing Attacks. PCWorld. Retrieved 19 May 2016, from http://www.pcworld.com/article/135293/article.html

[35] Phishing Techniques | Phishing.org. Phishing.org. Retrieved 19 May 2016, from http://www.phishing.org/phishing-techniques/

[36] Anti-phishing Working group, 2016. Retrieved 19 May 2016, from https://docs.apwg.org/reports/apwg_trends_report_q1_2016.pdf

[37] Badssl.com. Retrieved 27 December 2016, from https://badssl.com

[38] Parsons, K., McCormac, A., Pattinson, M., Butavicius, M., & Jerram, C. (2013, July). Phishing for the truth: A scenario-based experiment of users' behavioural response to emails. In IFIP International Information Security Conference (pp. 366-378). Springer Berlin Heidelberg.

[39] Sheng, S., Holbrook, M., Kumaraguru, P., Cranor, L. F., & Downs, J. (2010, April). Who falls for phish?: a demographic analysis of phishing susceptibility and effectiveness of interventions. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (pp. 373-382). ACM.

APPENDICES

APPENDIX A – LETTER OF APPROVAL

DALHOUSIE UNIVERSITY
Research Services

Social Sciences & Humanities Research Ethics Board
Letter of Approval

July 05, 2016

Manisha Arora
Computer Science\Computer Science

Dear Manisha,

REB #: 2016-3861
Project Title: Exploring user strategies in determining trustworthiness of websites and emails

Effective Date: July 05, 2016
Expiry Date: July 05, 2017

The Social Sciences & Humanities Research Ethics Board has reviewed your application for research involving humans and found the proposed research to be in accordance with the Tri-Council Policy Statement on *Ethical Conduct for Research Involving Humans.* This approval will be in effect for 12 months as indicated above. This approval is subject to the conditions listed below which constitute your on-going responsibilities with respect to the ethical conduct of this research.

Sincerely,

Dr. Karen Beazley, Chair

APPENDIX B – AMENDMENT APPROVAL

Social Sciences & Humanities Research Ethics Board
Amendment Approval

September 28, 2016

Manisha Arora
Computer Science\Computer Science

Dear Manisha,

REB #: 2016-3861
Project Title: Exploring user strategies in determining trustworthiness of websites and emails

The Social Sciences & Humanities Research Ethics Board has reviewed your amendment request and has approved this amendment request effective today, September 28, 2016.

Sincerely,

Dr. Karen Beazley, Chair

APPENDIX C – RECRUITMENT SCRIPT

We are recruiting participants to take part in a study to assess the trustworthiness of emails and websites. We are looking for participants who are 18 years or older, with at least 1 year of experience with the internet (i.e. be a regular internet user). We need participants with non-technical expertise.

Screening procedure will be conducted to finalize the participants for the study. Those interested in the study will be sent a screening questionnaire through Dal Opinio/online questionnaire to assess technical proficiency and suitability for the study.

The study will be conducted in the Graphics and Experiential Media (GEM) Lab, on the 4th floor of Mona Campbell building, Dalhousie University. You must bring your own laptop (Windows/Mac) to be used in the study. First of all you will meet the researcher, where the study will be explained in detail and you will be asked to provide informed consent. After this, the researcher will download a screen recording software on your laptop to collect data during the study and mention all study instructions. You will then complete a task to assess the trustworthiness of websites or emails which will be video recorded and followed by a post-observation questionnaire. After this, you will take part in a semi-structured interview which will be audio-recorded. The entire session will take about an hour to complete. Each participant will be compensated with $10 for participating in the study. Additionally, the top 25% of the participants in each category (technical and non-technical) will receive an extra $10 at the end of the research period.

If you are interested in participating, please contact Manisha Arora by email at Manisha.arora@dal.ca.

APPENDIX D – INFORMED CONSENT

Project title: Exploring user strategies in determining trustworthiness of websites and emails

Principal Investigator: Manisha Arora, a graduate student at Faculty of Computer Science,

manisha.arora@dal.ca

Contact Person: Manisha Arora, Faculty of Computer Science, Manisha.arora@dal.ca
Supervisors: Dr. Kirstie Hawkey, Faculty of Computer Science, hawkey@cs.dal.ca
Dr. Srinivas Sampalli, Faculty of Computer Science, srini@cs.dal.ca
Other team members: Dr. Raghav V. Sampangi, Postdoctoral fellow, Faculty of Computer Science Fatimah Alshammari, a graduate student at Faculty of Computer Science
Funding provided by: The study is funded by NSERC (Natural Sciences and Engineering Council of Canada)

Introduction:
We invite you to take part in a research study being conducted by Manisha Arora, a student at Dalhousie University as part of my computer science degree program. Your participation in this study is voluntary and you can withdraw from the study at any time. Your decision of taking part in the study or withdrawing will not impact your employment or performance evaluation. The study is described below in detail. It includes all the risks and benefits you might face during the study. Participation in the study may increase your browsing knowledge and we might learn things that will benefit others. You should discuss any questions you have about this study with Manisha Arora (Manisha.arora@dal.ca).

Purpose:
The purpose of our research is to explore the strategies used by you to assess the trustworthiness of websites while accessing websites from your own laptop using browser of your choice. In the beginning, researcher will explain the study in detail. After this, your browsing behavior will be observed. Then, you will complete a Post-observation questionnaire and finally you will be asked few questions in the interview. In the end of the session, you will be given a score sheet showing you the number of correct decisions made. The whole session will take approximately 1 hour in total.

Who Can Take Part in the Research Study?

The population for the study consists of internet users. You can participate in our study if you are a part of Dalhousie university faculty, students, and staff or are currently a resident of Halifax. This population will be composed of wide range of the technical and non-technical participants needed for the study. You should be 18 years of age or older, with at least 1 year of experience with using the internet.

What You Will Be Asked to Do:
Firstly, you will be asked to assess the trustworthiness of the websites or emails, i.e. whether these are legitimate or fraudulent and to explain the reasons leading you to make this decision. You will be seeing websites or emails by clicking on links provided on a

web interface/web page. During this part, you will be using your own laptop, talking out loud and a screen recording software will be installed on your machine with your permission for data collection purpose After the study, we will be sure to uninstall this software and collect the screen recorded data from your laptop. Your interaction with the websites or e-mails will be video recorded (Video cameras will be used to capture only your hand movements and where you are actually looking on the screen which will be used only for the purpose of analysis).

Secondly, you will be asked to complete a Post-observation questionnaire on Dal Opinio/Paper questionnaire and finally, you will take part in an interview, where you will be asked about your decisions regarding websites and emails in detail and this will be audio recorded.

Possible Benefits, Risks and Discomforts

Since you will be using your own laptop, there will be low risks associated with the study like those involved in normal internet surfing. There may be some risk of embarrassment for your wrong responses.

The direct benefit from the study will be increased browsing knowledge. As in the end of the session, you will be given a score sheet showing the number of correct or wrong decisions made by you during the observation. Participation in the study may increase your browsing knowledge and we might learn things that will benefit others. If you feel uncomfortable at any time of the study, you are free to stop participating and to withdraw the collected data up to that point in the study. You can withdraw the data until the end of your session.

You will be compensated with $10 for participating in the study even if you are not able to finish. As an incentive to give correct responses, the top 25% of the participants, as scored on their accuracy in detecting malicious and non-malicious emails/websites, will receive an additional $10 at the end of the research.

All the personal details will be kept confidential. Your comments may be quoted in the report and any published materials. If we include any of your quotes in our publications, you will be referred by a participation ID and not by your name. All the research data will be kept safely in a locked cabinet and the anonymity of the textual data will be preserved by using pseudonyms or participant numbers. The data will be kept in accordance with the Dalhousie university policy for five years.

If you have any ethical concerns about your participation in this research, you may also contact Research Ethics, Dalhousie University at (902) 494-1462, or email: ethics@dal.ca.

Signature Page

Project Title: Exploring user strategies in determining trustworthiness of websites and emails

Lead Researcher: Manisha Arora, a graduate student at Faculty of Computer Science, manisha.arora@dal.ca

"*I have read the explanation about this study. I have been given the opportunity to discuss it and my questions have been answered to my satisfaction. I hereby consent to take part in the study. However, I understand that my participation is voluntary and that I am free to withdraw from the study at any time."*

I understand that beyond my session it will be difficult to remove my data, as it may have been anonymized and compiled.

Participant	Researcher
Name: _____	Name: _____
Signature: _____	Signature: _____
Date: _____	Date: _____

"I agree that my participation in the experiments will be recorded for the purpose of analysis. Video recording for the observation phase, Screen recording for the observation phase and Audio recording for the semi-structured interview. I understand that this is a condition of participation in the study.

☐Yes ☐No

Participant	Researcher
Name: _____	Name: _____
Signature: _____	Signature: _____
Date: _____	*Date:* _____

"*I agree to let you directly quote any comments or statements made in any written reports, interview or observation phase and I understand that the anonymity of textual data will be preserved by using pseudonyms.*"

☐Yes ☐No

Participant Researcher

Name: _____ Name: _____

Signature: _____ Signature: _____

Date: _____ Date: _____

"If you are interested in getting a copy of publication, please check below and provide your email address."

☐ I would like to get notified via email when the publication is available
Email address for notification: _____

APPENDIX E: SCREENING QUESTIONNAIRE

1. **What is your age in years ?** _____

2. **Sex:**
   - Male
   - Female
   - Other

3. **What is the highest level of education that you have completed or are in the process of completion?**
   - High School
   - Community College (e.g. NSCC)
   - Under Graduate Degree
   - Master's Degree
   - Doctoral Degree
   - Other

   If other, please mention _____

4. **What is your field of study?(If relevant)** _____

5. **If working (part-time or full-time), what is your occupation?**

   _____

6. **Do you use anti-virus software on your computer? Yes _____No _____ Don't _____ know**
7. **How would you rate your technical expertise? I am the one who:**
   - Always helps others
   - Sometimes helps others
   - Does not seek others' help
   - Sometimes asks for help
   - Always asks for help

8. **How many years of experience do you have with the internet?**
   - None
   - One year
   - At least 1 but less than 3 years
   - At least 3 but less than 5 years
   - At least 5 but less than 10 years
   - More than 10 years
9. **How many years of experience do you have in computer security?**
   - None
   - One year

- At least 1 but less than 3 years
- At least 3 but less than 5 years
- At least 5 but less than 10 years
- More than 10 years

10. Which web browser do you use on daily basis for web surfing?

- **Google Chrome**
- **Mozilla Firefox**
- **Safari**
- **Internet Explorer**
- **Other**

If other, please mention ----------

11. Have you ever :

	Yes	No	Don't know
Designed a website?			
Registered a Domain Name?			
Used SSH?			
Configured a firewall?			

12. Are you currently using a Mac or PC?

- Mac
- PC
- Don't know

13. Please enter whether or not you have experienced any of the following:

	Yes	No	Don't know
Credit card fraud			
Stolen online password			
Stolen Social Security Number			
Identity theft			

14. Have you purchased something online in the past year?

- Yes
- No

15. Can you check your email from someone else's computer?

- Yes
- No
- Don't know

If yes, does it mean that you are not as good in managing security as you initially thought? Or, is it something else?

---

16. **Indicate how often do you use the following websites:**

Website	Never used	Use 1-10 times/year	Use 1-10 times/month	Use daily
Amazon.com				
Ebay.com				
PayPal.com				
Any Banking website				
Social Networking website				

APPENDIX F: POST-OBSERVATION QUESTIONNAIRE

1. What does https:// in the starting of a website address mean?

2. The lock icon present in the web browser indicates :

3. What does domain highlighting mean?

4. SSL/TLS certificate means :

5. Do you look at the URL bar to confirm that you are visiting an intended website?

6. Do you visit websites you have not heard of before?

7. When you click on a link in an email and that link takes you to a website that asks for your password, do you enter it?
    - Yes, often
    - Yes, rarely
    - No
    - I don't know

8. Do you click on links in an email that are sent by unknown people or companies to you?
    - Yes, often
    - Yes, rarely
    - No
    - I don't know

9. Do you respond to any email that requests you to submit your personal information (such as credit card details, etc.) and is sent by people or companies you don't know?
    - Yes, often
    - Yes, rarely
    - No
    - I don't know

10. Do you consider the emails with bad grammar or spelling mistakes as suspicious?
    - Yes, often
    - Yes, rarely
    - No
    - I don't know

APPENDIX G: SEMI-STRUCTURED INTERVIEW GUIDE

1. What do you understand by the term phishing?
2. Do you visit websites you have not heard of before? Why or why not?
3. Did you check the presence of HTTPS in the starting of website address? Why/why not?
4. Did you look at the URL bar to confirm that you are visiting an intended website? Why/why not?
5. Did you check the lock icons in the browser? What does they indicate?
6. Were you suspicious of links received in the emails? What made you suspicious?
7. What do you understand by SSL/TLS certificates? Do you check the SSL/TLS certificate while visiting a website?
8. What features of a website you consider important for the identification of phishing attacks?
9. How much did the following factors influence your decision to assess the website as legitimate or fraudulent? Rate from 1(no influence) to 5 (strongly influence).

Website Features	**Your ratings**				
	1	**2**	**3**	**4**	**5**
Content					
URL					
Look and Feel					
Other Factors					

Please mention the other factors:

10. What do you think about the effectiveness of browsers' security cues? Please suggest some improvements.
11. What are the things you keep in mind in order to protect yourself from phishing websites and e-mails?
12. Are you aware of any phone phishing scams? How do you judge it as legitimate or not?

APPENDIX H: CODING SHEETS

Operating System	**Mac**	
	Windows	
Browser	**Chrome**	
	Firefox	
	Safari	
	Internet Explorer	

Website	Actual Type	Answered Legitimate	Answered Phishing
1	**0**		
2	**0**		
3	**0**		
4	**1**		
5	**0**		
6	**1**		
7	**0**		
8	**1**		
9	**0**		
10	**0**		
11	**0**		
12	**1**		
13	**0**		
14	**1**		
15	**0**		
16	**1**		
17	**0**		
18	**0**		
19	**1**		

(0: Phish, 1: Legi)

Websites	Important information in decision making	Proceed?
1		
2		
3		
4		
5		
6		
7		
8		
9		
10		
11		
12		
13		
14		
15		

16		
17		
18		
19		

Email	Actual Type	Answered Legitimate	Answered Phishing
1	**0**		
2	**0**		
3	**0**		
4	**0**		
5	**1**		
6	**0**		
7	**1**		
8	**0**		
9	**1**		

Email	Important information	Proceed? (link)	Website			Proceed?
			Type	Ans Phish	Ans Legit	
1			0			
2			0			
3			0			
4			0			
5			0			
6			0			

7			1			
8			0			
9			1			