END USER MENTAL MODELS OF SOCIAL ENGINEERING ATTACKS

by
Lin Kyi

A thesis submitted to
the Faculty of Graduate and Postdoctoral Affairs
in partial fulfillment of
the requirements for the degree of

MASTER OF ARTS

Human Computer Interaction

at

CARLETON UNIVERSITY

Ottawa, Ontario
August, 2021

Abstract

How do end users understand social engineering attacks, and how do their mental models differ from reality? To investigate, we have proposed a new social engineering attack framework, and ran two studies using the framework as the foundation. In the first study, we conducted 30 interviews to investigate social engineering mental models, and found that *confidence* and *accuracy* are underlying themes that affect users' mental models. In the second survey, we quantified how confidence and accuracy impact mental models at different stages of an attack. We found that users tend to be overconfident in their ability to understand social engineering attacks, but hold inaccurate beliefs. They hold major misconceptions of what constitutes as social engineering, and the threat levels of these attacks. Based on our results, we have proposed various educational and design opportunities to match social engineering mitigation strategies to end user mental models of social engineering.

Acknowledgements

I would like to thank my family and friends for supporting me throughout my thesis, and for also acting as pilot participants. Thank you to Dr. Robert Biddle for providing me with insightful feedback for my thesis.

I also would like to give a big, big thank-you to Dr. Elizabeth Stobert, who has supported me at every step of my thesis, and for teaching me very valuable research skills. I really appreciate all of the support and attention you give all of your students, and am always amazed at how you balance everything so well!

Table of Contents

Abstract ii

Acknowledgements iii

List of Tables vii

List of Figures viii

Chapter 1 Introduction 1

1.1 Research Problem 1

1.2 Contributions 2

1.3 Thesis Outline 3

Chapter 2 Background 4

2.1 Introduction 4

2.2 Prevalence 5

2.3 Social Engineering Mitigation Strategies 6

2.3.1 Technical Strategies 6

2.3.2 Educational Strategies 9

2.4 How Users Understand Social Engineering 11

2.5 Mental Models of Social Engineering Attacks 12

2.6 Social Engineering Attack Frameworks 13

2.6.1 Existing Social Engineering Frameworks 14

2.7 Research Problem 16

Chapter 3 Proposed Social Engineering Framework 17

3.1 Attack Formulation 18

3.1.1 Attack Vector 18

3.1.2 Audience 18

3.2 Persuasive Techniques 19
3.2.1 Types of Persuasive Techniques 19
3.3 Call to Action 21
3.4 Utilize the Information 22
3.5 Expands Scale 22

Chapter 4 User Mental Models of Social Engineering Attacks 23
4.1 Description 23
4.2 Methodology 23
4.3 Participants 24
4.4 Analysis 25
4.5 Results: Open Coding 25
4.6 Results: Themes 29
4.6.1 Stage 1: Attack Formulation 30
4.6.2 Stage 2: Persuasive Techniques 35
4.6.3 Stage 3: Call to Action 40
4.6.4 Stage 4: Utilize the Information 43
4.6.5 Stage 5: Expands Scale 45
4.7 Mental Models of Social Engineering Attacks 46
4.7.1 Confidence and Accuracy 47

Chapter 5 Examining Users' Confidence and Accuracy by Attack Stage 48
5.1 Description 48
5.2 Methodology 48
5.3 Participants 51
5.4 Hypotheses 51
5.5 Results 54
5.5.1 Confidence 54
5.5.2 Accuracy 55
5.5.3 Confidence and Accuracy Interactions 56

5.5.4 Exploratory Results 59

Chapter 6 Discussion **62**

6.1 Design and Education Recommendations 65

Chapter 7 Conclusion **67**

7.1 Limitations 67

7.2 Future Work 68

Bibliography **69**

Appendix A Appendix **75**

A.1 User Study 1 75

A.1.1 Demographics Questionnaire 75

A.1.2 Interview Questions 77

A.1.3 Ethics Application: Study 1 77

A.2 Study 2 97

A.2.1 Survey 97

A.2.2 Table of Attack Examples 102

A.2.3 Ethics Application: Study 2 105

List of Tables

2.1 Social engineering attack examples. 4

4.1 Participant demographics 25
4.2 Codes used in the open coding process. 26

5.1 Participant demographics 52
5.2 Confidence results 54
5.3 Accuracy results 55
5.4 Probability of accuracy: Chi-squared results 56
5.5 Wilcoxon results by the stages of the framework 57
5.6 Wilcoxon Results by Attack Vector 60

A.1 Description of the social engineering attack examples used. . . 102

List of Figures

2.1 An image of Google Chrome's current lock icon. 7
2.2 An example of the PhishGuru embedded training given to some participants [36]. 10
2.3 An image from the Anti-Phishing Phil game. 11
2.4 Mitnick's original social engineering framework, illustrated by Mouton et al. [42]. 14
2.5 Mouton et al.'s expanded framework [42] 15

3.1 The proposed social engineering attack framework. 17

4.1 The manual theme-building process. 30

5.1 Screenshot of an SMS social engineering attack we showed participants. This example is a generalized attack trying to elicit fear to gain money from users. 49
5.2 Screenshot of a social media social engineering attack we showed participants. This example is a targeted attack trying to appeal to greed to gain money. 50
5.3 Accuracy rates by the stages of the framework. 55
5.4 Distributions of aggregated confidence and accuracy scores by the stages of the framework. 58
5.5 Distributions of aggregated confidence and accuracy scores by attack vector. 61

Chapter 1

Introduction

1.1 Research Problem

Social engineering, which refers to attacks that manipulate users into revealing information to compromise computer systems [34, 41], is one of the most commonly reported kinds of internet fraud [43]. Due to its prevalence, social engineering awareness amongst users is increasing, yet many users and organizations are still vulnerable [15,20,33,34]. In 2017, 29 billion spam emails were sent daily; 1 in every 329 users fell for a phishing email [11]. Large companies are not immune to social engineering attacks either; 75% of organizations have been victim to spear phishing attacks [6].

A lot of social engineering research in usable security pertains to cues users look at to determine legitimacy [18,33,62], vulnerability to phishing [20,50], and evaluations of phishing solutions [38,51]. As a result, little is known about users' behaviours and mental models at different stages during a social engineering attack. We believe that to effectively protect users from social engineering attacks, we must understand their mental models of the different stages within the social engineering attack cycle to identify where exactly users need more education and technical help. We must understand where users' mental models differ from reality, and address these differences in order to better protect them.

In this thesis, we ran two studies, and used a mixed methods approach to investigate end user mental models of social engineering attacks and to understand where users are performing poorly when it comes to social engineering attacks. In the first study, we conducted semi-structured interviews with users to investigate their mental models of the different stages within a social engineering attack. Using thematic analysis, we found that the underlying themes of *confidence* and *accuracy* were impacting user understandings of social engineering attacks. In the second study, we conducted a survey to investigate where in the social engineering attack cycle users

are performing worst by measuring participants' confidence and accuracy. We found that participants' accuracy scores were relatively low, and were most likely the result of guessing. However, confidence scores were consistently high, regardless of the stage. This overconfidence in inaccurate beliefs is especially dangerous, especially as participants had misconceptions of social engineering, such as mistaking fraud, spam, and online ads as attacks, and believing that social engineering attacks are not as threatening because participants tend to view social engineering attacks as a nuisance, rather than a real threat. This suggests that users have flawed mental models of social engineering, revealing future opportunities for user interface designs and education.

1.2 Contributions

The contributions of this thesis are as follows:

- We break down the social engineering attack cycle into different steps, and understand user mental models of these steps so we are better informed about where exactly users hold incorrect mental models, and where they are opening themselves up to risk. To the author's knowledge, there are no studies looking at user mental models of the whole social engineering attack cycle.

- This study is one of the first to look at mental models of other social engineering attack vectors, such as SMS and social media phishing. Most research in social engineering tend to focus on email [50] and website [18, 38] phishing, so we expand the research about how users understand and behave with these emerging attack vectors.

- Study 1 uses qualitative methods to investigate mental models of social engineering attacks, and identified how users understand different aspects of the social engineering attack cycle.

- Study 2 quantifies how confidence and accuracy play a role in social engineering mental models, and identified which stages participants especially struggled with.

1.3 Thesis Outline

This thesis is organized as follows:

- Chapter 2 reviews the previous research, including the various kinds of social engineering attacks, current mitigation strategies, and previous research on social engineering mental models.
- Chapter 3 is where we propose our updated social engineering attack framework.
- Chapter 4 summarizes the methodology and results of the first study which focuses on user mental models of the social engineering attack cycle.
- Chapter 5 summarizes the methodology and results of the second study which focuses on investigating where in the attack cycle users tend to have the most trouble.
- Chapter 6 discusses the implications of the two studies, and the broader applications of this research.
- Chapter 7 presents our conclusion, and includes the limitations of the studies, and future work.

Chapter 2

Background

2.1 Introduction

Social engineering in IT security refers to attacks where users are manipulated into compromising computer systems by revealing confidential information, or installing malware, etc. through various psychological manipulations [34,41]. Social engineering attacks are effective because most users are poor at detecting them [15,20,33,34], and technical controls often do not account for user behaviours [20, 34]. In essence, it is often easier to trick humans into compromising information than it is to break into a security system [41, 60].

Within social engineering exist many types of attacks. Phishing, especially email phishing and phishing websites are the best-known and most commonly used type of social engineering attack [2, 5, 60], but there are many other emerging social engineering attacks which are less well-known by end users, and seldom studied by researchers [25, 34]. Table 2.1 shows a description of the various social engineering attacks. This is not an exhaustive list, and there are several other types of social engineering attacks, but for this project we are mainly focusing on remote attacks such as those listed in Table 2.1. Remote attacks are attacks where the attacker does not need to directly interact with the target to conduct the attack, and can be accomplished without in-person interactions.

Table 2.1: Examples of common social engineering attacks.

Social Engineering Attack Name	**Description**
Email phishing	An email, appearing to be from a legitimate source, will request a user to click a link or download a file [20].

Continued on next page

Table 2.1 – *Continued from previous page*

Social Engineering Attack Name	**Description**
Water-holing	A website a target is expected to go to is infected with malware; attackers look for vulnerable websites and vulnerabilities in the security system to exploit [34].
Vishing	VOIP technology is used to create phishing calls to attain personal information [25, 44].
QRishing	Phishing through QR codes; users scan an unknown QR code, and are led to the phishing item, such as a website or downloading an item [57].
Fraudulent apps	An app is coded to look like a similar legitimate app, or a vulnerable app is given malicious code.
Malvertisement	Using malicious ads to install malware on a computer [28]
WiFi evil twin	A spoofed WiFi access point that appears to be legitimate to steal user information [28].
SMiShing	Using SMS to send phishing links and files to users [30].
Baiting	The attacker offers a phishing item disguised as something appealing to the target [34].
Trojan Horse	Malware that is disguised as a legitimate software [28].
Drive-by download	Placing a malicious file into a website for users to download, not knowing it is malware [28].
Scareware	Scare users into downloading or buying unnecessary software for a fake virus [28].

2.2 Prevalence

Social engineering is very common [14]; in 2017, over 29 billion spam emails, many of which were phishing emails, were sent each day [11]. According to the 2020 Verizon Data Breach report [2], 22% of data breaches begin through phishing. Around 75% of organizations, meaning attacks which target employees of an organization, have been victim to attempted spear phishing attacks [27], and each successful attack on an organization results in a $25 000 to $100 000 gain. On average, fraud costs end users between $218 to $364 per successful scheme [14].

After hacking, phishing is most commonly used for data breaches, and phishing is the most common social engineering attack [2, 14]. Despite phishing being relatively well-known, it is still a threat to organizations and end users alike [14, 20, 39]. Email phishing is the most common social engineering vector, but social engineering websites are becoming more common [5]. According to Google's 2020 Safe Browsing List and the APWG's June 2020 report, thousands of phishing websites exist on the Internet [2, 3], and many of them are returned in Google search results [3].

Social engineering still continues to be a problem, despite technical controls and user education efforts [14, 20, 50]. The problem will continue to get worse with new attack vectors emerging [20, 24]. Phishing and social engineering attacks have increased in the past couple of decades because of an increase in mobile phone use [30]. Mobile phones provide more attack vectors, such as social engineering through SMS, QR codes, social media, and mobile apps, and these emerging attacks are less well-known, which will make many users vulnerable [24, 25].

However, we also note that finding reliable prevalence statistics is inherently difficult. Most of the statistics we found came from organizations, and their security breaches [2, 14, 27]. It is unclear exactly how many users have fallen for social engineering attacks, and which ones they fell for. These prevalence statistics for end users are difficult to collect because users are not always aware if they fell for a social engineering attack [46, 49, 61]. Users may also feel embarrassment that they fell for an attack [32], and may not report it out of shame, or fear of employer discipline.

2.3 Social Engineering Mitigation Strategies

There are a variety of technical and educational strategies to help prevent users from falling victim to social engineering attacks. However, these mitigation strategies can only do so much; they do not protect users from all attacks, and it is a constant competition with attackers to come up with new mitigation strategies [20].

2.3.1 Technical Strategies

Technical mitigation strategies attempt to help users in determining whether an item is legitimate or fraudulent. Since many users are not very good at accurately detecting

social engineering attacks [15, 20, 33, 34], technical strategies attempt to reduce the guesswork done by users, and provide a guide to users when they encounter suspicious items.

A commonly used social engineering mitigation strategy for websites is the use of browser indicators, such as a lock or shield icon, to indicate a website's security state [23]. For example, Google Chrome will display a lock icon if a website is secure (Figure 2.1). Most participants look at the lock icon of a web page, even if they do not understand what it means [62], and most users associate HTTPS with a secure internet connection [23]. HTTPS is more secure than HTTP because it encrypts any data that is exchanged between a user's browser and the website they are visiting, therefore adding another layer of security [2]. Google Safe Browsing will also alert users of potential phishing and malware sites, and will block these sites for the user. However, the majority of browser cues have been developed for computers with larger displays; mobile browsers are harder to design for because they have a smaller screen [7]. Therefore, the smaller screen lacks some of the browser indicators seen on larger screens [7, 24].

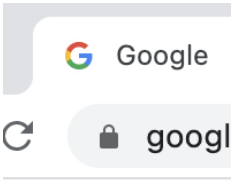


Figure 2.1: An image of Google Chrome's current lock icon.

Many technical controls are not effective because they do not account for human behaviour [6, 34]. Technical controls are either not understood by users [18, 23], or ignored [18]. In Dhamija et al.'s 2006 study, anti-phishing cues were not noticed by 23% of participants; users did not look at a website's URL, status bar, or other security indicators [18]. Instead, these users relied on the website's look and feel to determine legitimacy [18]. Pop-up warnings for fake website certificates were ignored by 15 out of 22 participants in this study [18]. In Felt et al.'s 2016 study on Chrome browser indicators, it was found that most users did not understand what different security indicator icons meant [23].

To make technical controls more effective, it is important that social engineering

mitigation strategies align with user mental models and user behaviours to ensure they are effective [33]. Browser indicator icons which are well-designed, and clearly convey to the user what they do are more effective [23]. Anti-phishing tools which interrupt the user's task are more effective, and ideally, should be within the browser so users do not need to install any software [22]. Users are more likely to follow anti-phishing cues if they are relevant to the current website they are using, and if the cue uses language they understand [33].

Given that there is still user error when it comes to social engineering mitigation strategies, some solutions are automatically implemented. A common method of preventing social engineering attacks is through blocklisting. This is where suspicious IP addresses, URLs, and keywords are collected, and considered when identifying whether an item is legitimate or not [24]. Blocklisting is implemented on several devices and platforms, such as Google Chrome and Gmail [24]. Blocklisting is effective, but only when it is regularly updated [52]. For SMiShing attacks, Yadav et al. proposed SMSAssassin to filter out spam messages. Using a crowd-sourced list of words and message patterns commonly used in SMiShing, spam messages are identified. SMSAssassin has a 72.5% accuracy rate for identifying spam messages [66].

Unfortunately, technical strategies may not only misunderstood by users, but they can also be incorrect. Recently, there has been an increase in the number of phishing websites that bypass security warnings because they have an HTTPS web address, and a valid website certificate [2]. This is partly due to the *Let's Encrypt* movement, which started as an effort to make every website secure by automatically providing free web certificates [1]. The movement to secure websites with HTTPS and provide legitimate web certificates has prevented unwanted parties from snooping and taking advantage of vulnerable websites. However, the goal of making it easier for websites to get web certificates has also provided phishing websites with legitimate certificates [56]. This makes phishing websites appear more legitimate because these websites, which were originally difficult to get a web certificate for now have one, and have the secure lock icons associated with a secure website. In a recent APWG report, around 89% of phishing websites in 2020 had a Domain Valid (DV) web certificate, which is often given for free, and is the weakest form of validation [2, 56].

Password managers can help prevent users from giving login credentials to illegitimate websites [53]. Although there is inconsistency between the different password managers available, many of them do not fill in passwords for websites where the domain name does not match the original website the password was saved to [53]. This means that for phishing websites that are made to look like a legitimate website, password managers will often not fill in user credentials, therefore protecting the user from a potential attack.

From these technical strategies, it is evident that preventing users from falling victim to social engineering attacks is a multifaceted problem, and technical solutions are not always perfect. In addition to technical mitigation strategies, the user also plays a part in social engineering attack mitigation.

2.3.2 Educational Strategies

A common method to prevent users from falling victim to social engineering is by educating them. End users are not often computer security experts, therefore they need to be educated about how to make secure decisions. This can come in the form of websites, flyers, posters, and computer literacy training [60]. User security education and training is the most commonly used method to change behaviours [60]. Research has shown that education can improve email phishing detection, if done effectively [38]; in one study, education created a 40% reduction in users giving important information to attackers in one study by Sheng et al. [50]. Educational strategies have often been applied to reduce the likelihood of falling victim to email phishing.

Embedded training, which is where users who have fallen for a phishing attack are then presented with phishing training, has been shown to be an effective method of teaching users by motivating them [37]. Due to its effectiveness, embedded training has been implemented in many commercial phishing prevention solutions, such as Proofpoint, and CoFense [60].

Kumaraguru et al. created PhishGuru, an embedded phishing training system which tries to educate users to prevent falling victim to phishing attacks [36]. Figure 2.2 shows an image of one of the embedded training messages given to participants in one condition. In their study with 515 participants, it was found that PhishGuru

was effective for helping users retain phishing knowledge one month after they were trained, a second training message was effective for preventing users from giving phishing websites information, and that this embedded training does not prevent users from clicking links in legitimate emails [36].

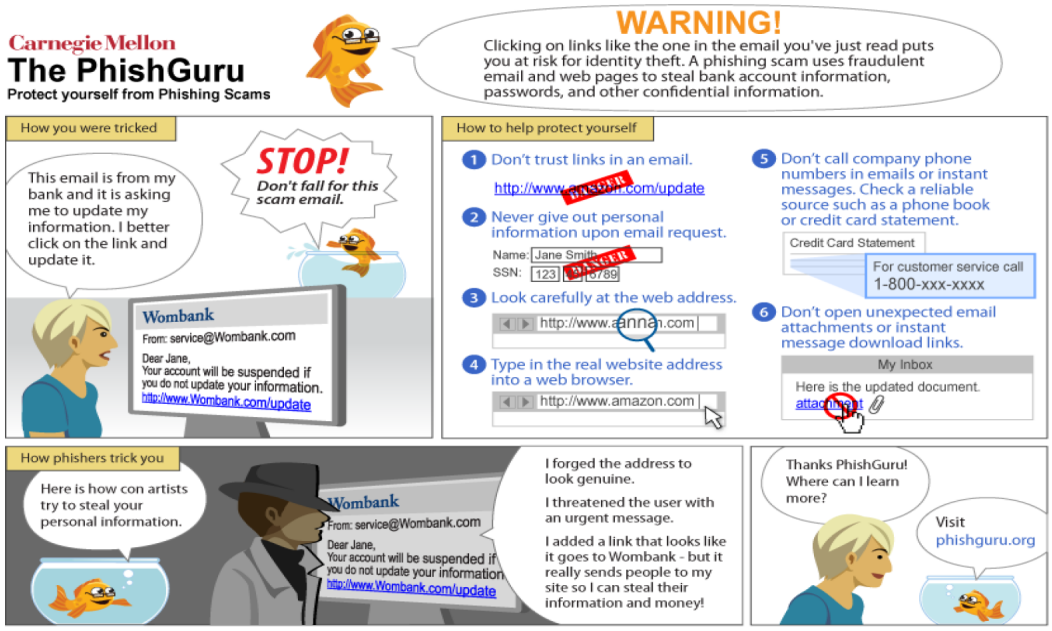


Figure 2.2: An example of the PhishGuru embedded training given to some participants [36].

In addition to the actual training materials, education should also focus on how it is conveyed. A study by Wash and Cooper found that a user's perceived origin of phishing training materials impact security decisions. The method of giving only facts and security advice worked when it was told to users by security experts, and stories of security mishaps worked to improve phishing awareness only when told by a peer [60]. Social engineering mitigation strategies should also dispel common myths; users should be shown how easy it is for attackers to mimic websites [47].

An interactive and engaging way of educating users is through online games. Sheng et al. created *Anti-Phishing Phil* (Figure 2.3), a game which teaches users to decrease their phishing susceptibility [51]. They found that participants who played the training game were significantly more effective at identifying phishing URLs compared to reading printouts with security advice and existing anti-phishing tutorials [51].

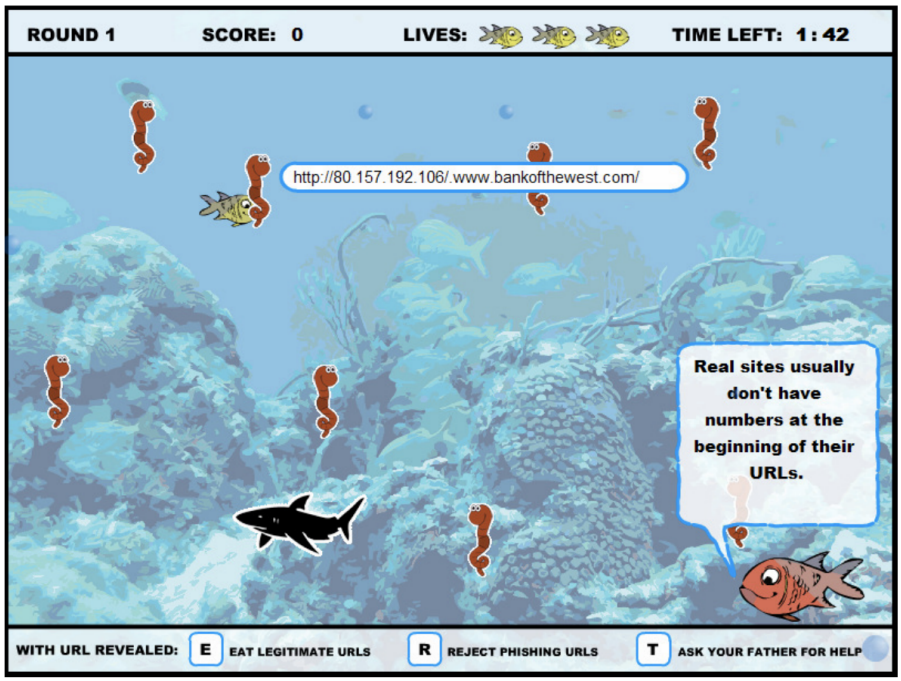


Figure 2.3: An image from the Anti-Phishing Phil game.

2.4 How Users Understand Social Engineering

Users are generally poor at recognizing social engineering attacks, and understanding what to do when they encounter one [18, 59, 68]. Users remain susceptible to social engineering, despite mitigation strategies being implemented. This is due to three factors. Firstly, users lack knowledge about computer security and social engineering; users do not understand security indicators, and hold incorrect beliefs about cues [18]. Secondly, attackers deceive victims by mimicking legitimate methods [18], such as making a phishing website look like the authentic one, or using HTTPS encryption for a phishing website [2]. Lastly, users fall for social engineering because security is not a primary task for them; users are not as focused on paying attention to security indicators because they are focused on other things [18].

Approximately one in every 392 users fall for phishing [11]. In a study with employees of a company where data security was a important, when employees clicked phishing links, 80% of them gave their username, and 60% gave their password [45]. Additionally, in an online survey on phishing, 32% of phishing emails were wrongly classified, and 26% of participants were unsure whether an email was a phishing email

or not [31]. Users do not look at the correct cues to judge if a site is legitimate or not. Users tend to base the legitimacy of a website on its "look and feel" rather than considering the URL, HTTPS encryption, or looking at the SSL certificate [18]. In a study on government employees, it was shown that 40% of employees have no strategy to recover their compromised accounts [40].

Education and user awareness for detecting social engineering are not comprehensive enough, and not fast enough to match attackers' methods. Attackers are improving their methods at an alarming rate which user awareness and technical controls cannot meet [2]. In 2020, 89% of phishing websites used HTTPS encryption to trick users into thinking their website is legitimate [2]. In comparison, less than 10% of phishing websites used HTTPS encryption in 2016 [2]. Attackers are learning what has been taught to users, and are now using it for their own ends.

2.5 Mental Models of Social Engineering Attacks

A mental model is defined as an internal representation of a concept, and it develops through experience and knowledge [49]. In this thesis, we will not be investigating how participants' mental models of social engineering are formed, nor how they are impacted by experience and reasoning. Instead, we will investigate the overall mental models of social engineering, which have already been influenced by a users' previous experiences.

Mental models have been shown to inform user security behaviours [59], therefore it is important to consider end user mental models when designing security systems. An incorrect mental model can result in insecure behaviours and a mistrust in technology [12]. Generally, users behave more securely if their mental model includes some understanding that they are vulnerable to security risks [59].

There are three groups users may be in when it comes to their mental models. The first group is where a topic is new to them; this group needs the topic addressed at their level in order to be effective to the user [63]. The second group is where users have a correct but limited mental model of a topic. To address this user group, one needs to understand what the user knows, and educate them by connecting new content to their pre-existing knowledge [55]. The last mental model group includes

those with an incorrect mental model. This group will need to be shown where and why they are wrong, and be re-educated [9].

Previous research has shown that generally, end user mental models of phishing and social engineering are flawed [10,20,68]. Users perform well when recognizing familiar scams, but are not as successful at identifying unfamiliar scams [20]. Cues such as the presence of an HTTPS URL, secure lock icon, and spoofing produced different levels of user awareness when it comes to recognizing fraudulent content [18,20]. When it comes to phishing, feedback is not immediate, and those who were phished may not realize until much later, or never realize where they were vulnerable [46, 49, 61].

In a study looking at how expert and novice mental models of phishing compared, it was found that experts have a more complex understanding of phishing compared to novices [68]. The perceptions end users have of phishing also influence the actions they took to protect themselves, if at all [68]. In another study looking at novice and expert security behaviour, it was shown that novices consider the security risks after they engage with an item, whereas experts consider the risks before [10]. Novices also tend to judge the legitimacy of a website or email by the way it looks [10], without understanding that the look and feel can be easily copied.

The majority of studies looked at specific aspects of phishing, which leaves other social engineering attacks understudied. Additionally, there is a lack of cohesion about mental models of social engineering as a whole. It is important to understand user mental models of social engineering from a broader perspective to properly address user needs and gaps in understanding, and to identify where users struggle.

2.6 Social Engineering Attack Frameworks

Social engineering frameworks allow security researchers to compare and understand social engineering attacks in a standardized method [42]. Additionally, a social engineering attack framework is useful because it allows us to better understand these attacks, and how they work.

To our knowledge, there are two previous social engineering frameworks; the social engineering framework was originally proposed by Mitnick in the early 2000s [41], and Mouton et al. built upon this model in 2014 [42]. Since these frameworks have

been proposed, new social engineering attack vectors and trends have emerged [2,14]. These frameworks are relatively outdated, and cannot be applied to all types of social engineering attacks, especially for attacks conducted remotely. An updated framework is needed to encompass emerging social engineering attacks, and address other psychological vulnerabilities being targeted.

2.6.1 Existing Social Engineering Frameworks

In Mitnick's original social engineering framework proposed in *The Art of Deception* [41], there are four steps (Figure 2.4). First is the *Research* stage. The attacker obtains information about their target; the more information known about the target, the better. The second step is *Developing Rapport and Trust*, which is where the attacker builds trust with the victim, who will be more willing to comply if they trust the attacker. Next, the attacker needs to *Exploit Trust*; they will request the victim give them information or do a task which will give the attacker access to the victim's information. The fourth step is to *Utilize the Information* gained from the attack. Mitnick states that this cycle may need to be repeated until the attacker achieves their target goal [41].

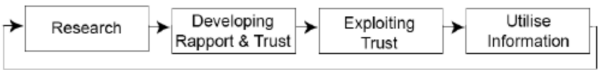


Figure 2.4: Mitnick's original social engineering framework, illustrated by Mouton et al. [42].

Mouton et al. [42] expanded upon Mitnick's original framework in 2014, since the original model was vague [42] (Figure 2.5). In the expanded model, the first step is the *Attack Formulation*. The attacker needs to identify their goal, and their target of the social engineering attack. Next, the attacker needs to *Gather Information* about the victim to develop a relationship. The third step is to *Prepare for the Attack*; the attacker ensures everything is ready, and develops an attack vector. The fourth step is to *Develop a Relationship*; the attacker has to establish trust with the victim and build rapport. After establishing a relationship, the attacker has to *Exploit the Relationship*. The attacker ideally will manipulate their target into a desired psychological state, then attain their desired information from the target. Last is

the *Debrief* stage, where the attacker ensures the target does not recognize they were attacked. If the goal is not reached, the attacker needs to go back to the *Information Gathering* stage.

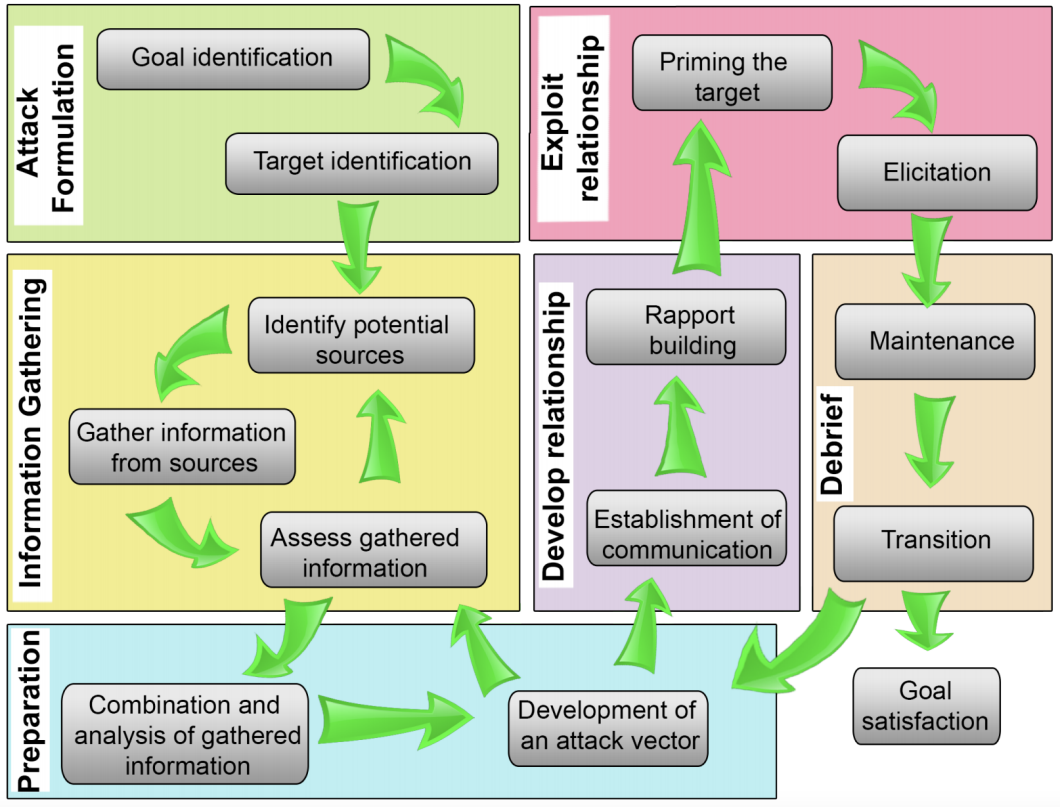


Figure 2.5: Mouton et al.'s expanded framework [42]

Problems with the Existing Frameworks

The existing models apply better to attacks that are individualized, where there are fewer targets. Most modern social engineering attacks are not like this, and target several users at once [11].

Mitnick's and Mouton et al.'s frameworks state that the attacker should thoroughly research their target, and imply that there is only one (or very few) targets. In modern social engineering attacks, while spear phishing is common, it is more common for attacks to be sent out to several individuals or organizations at once [11]. Researching targets may be a difficult and unrealistic task for modern social engineering attacks, which may use updated technologies to target attacks to a wider group

of individuals.

The existing frameworks also focus closely on building rapport and trust, which is not always the case in all social engineering attacks. In some social engineering attacks such as spear phishing, building rapport is important, but this is not the case for other remote attacks, such as drive-by download attacks or generalized phishing emails, where users are deceived through other means that do not involve fostering a relationship. We should recognize that other psychological techniques can be manipulated by attackers. Additionally, Mouton et al.'s expanded framework mainly focused on how Cialdini's Principles of Persuasion are exploited, but other psychological vulnerabilities such as fear and greed can also be targeted [33, 57].

The previous frameworks focused on the target giving information or doing something which gave the attacker access to their system. However, in social engineering attacks, we believe there is instead a tangible interaction the target engages in, such as clicking a link or downloading a file, which exposes the target to the attack. The previous frameworks are vague about exactly how the target exposes themselves to a social engineering attack, and focuses more on direct interaction with the target to gain access to their system. This is not the case for many modern social engineering attacks, where users are not directly giving information (whether they know or not) to an attacker, and instead is done through indirect means, such as by impersonating legitimate entities and having users give personal information to impersonated accounts, websites, etc.

2.7 Research Problem

We believe that social engineering mitigation strategies could be more effective if they also account for end user mental models of social engineering attacks. Previous studies have not examined user mental models, leaving a gap in our understanding of how to better protect users from social engineering.

By understanding user mental models based on our proposed social engineering framework, we believe we can work towards more effective social engineering mitigation strategies that align with not only user mental models of social engineering attacks, but also how social engineering attacks generally work.

Chapter 3

Proposed Social Engineering Framework

Drawing upon Mitnick and Mouton et al.'s frameworks, we propose a social engineering framework that is more suitable for a broader range of modern social engineering attacks. Our proposed framework, shown in Figure 3.1, is composed of five steps. The first step is the *Attack Formulation* stage; the attack vector, and the audience is identified here. Second, *Persuasive Techniques* are used to exploit the psychological vulnerabilities of users. The third step is a *Call to Action*; there must be a tangible interaction the user directly interacts with, such as by clicking a link or downloading a file to initiate the actual compromise. Fourth, the attacker then *Utilizes the Information* to work towards their target goal. Lastly, the attack *Expands Scale*; this is an optional step where the attack spreads to other individuals, or is repeated.

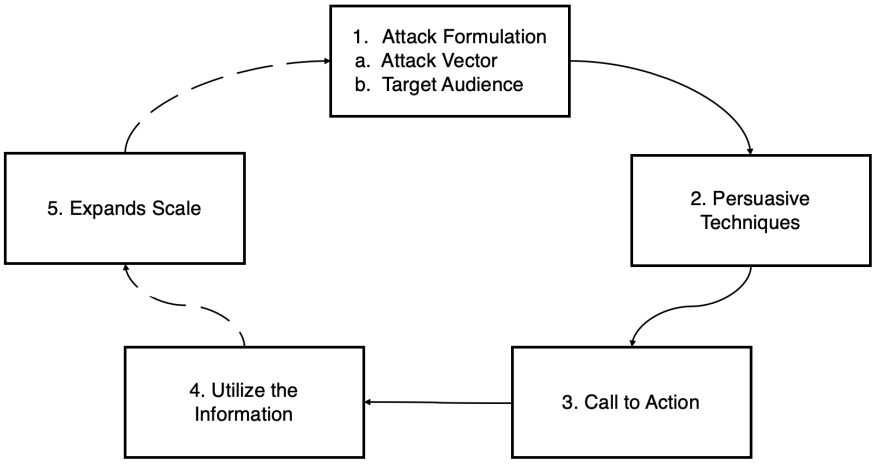


Figure 3.1: The proposed social engineering attack framework.

3.1 Attack Formulation

In this step, the social engineer plans the attack. They must identify their attack vector, what they hope to achieve, and their target audience, in no particular order.

3.1.1 Attack Vector

There are several methods an attacker may use to conduct their attack. The attacker may decide whether they want to do a physical social engineering attack (e.g., baiting targets with USB keys), or reach their victims remotely. Email is very popular [5,25], but attacks may also be conducted using SMS, social media, websites, QR codes, advertisements, and pop-ups.

3.1.2 Audience

Social engineering attacks also include a target audience, such as an individual, groups of individuals, or organizations. Sometimes, an attack may have no specific audience, but these are often less successful [29]. Target groups can be anywhere from general, to very specific groups or individuals. An attacker may target the elderly [44], younger individuals [50], users of social media [58], those who have smartphones [57], newcomers to a country or geographic area, speakers of a certain language, or users of certain websites or products so that the attack message is more specific, therefore more persuasive.

Attack scalability, referring to how many people are targeted in the attack and how/if the attack spreads to other users or devices, needs to be considered in this stage. The breadth of the attack audience will impact the future of an attack, such as the persuasive techniques, goals, and attack vector.

Targeting the desired audience with the correct medium is important, as the demographics of technology use are varied [13, 21, 44]. For example, vishing is often more successful when it targets elderly individuals, as they are more likely to trust phone calls, and are easier to reach via telephone [44].

3.2 Persuasive Techniques

The types of persuasive techniques used by attackers to fool their victims can vary. They can range from deception, to more advanced psychological techniques. Social engineering attacks are one of the most effective kinds of attacks because they target human weaknesses by exploiting human psychology [34,43]. The previous framework mainly focused on Cialdini's Principles of Persuasion (reciprocity, consistency, social proof, liking, authority, and scarcity) [16], but we propose a more extensive range of persuasive techniques because research has shown that social engineers successfully exploit other psychological vulnerabilities outside of these persuasive techniques [33, 57].

3.2.1 Types of Persuasive Techniques

Cialdini's 6 Principles of Persuasion

are the classic Persuasive Techniques which are exploited in social engineering attacks [41, 42]. The persuasion principles are:

- Reciprocity: Give the target something they may want, and they will feel the need to reciprocate [16].
- Consistency: Users like being consistent with what they have previously done. Once something has been done, they will want to follow through with it [16].
- Social Proof: When unsure what to do, users will tend to follow what those around them are doing [16].
- Liking: Users prefer to comply with those who they like [16].
- Authority: Users are more likely to comply with those who seem to have knowledge or authority [16].
- Scarcity: When something seems scarce, users will want it [16].

Previous research has shown that liking, scarcity, social proof, and authority are especially effective in SE [6].

Pretexting:

refers to giving the target a fake background story to elicit an emotion, such as empathy or fear. Once the background story is given, the attacker will follow up with the attack [26]. Since pretexting is not the attack in itself, but is rather being used to leverage psychological vulnerabilities to conduct the attack, we classify it as a persuasive technique.

Impersonation and pretexting are often classified as social engineering attacks in previous literature [8,67]. In this paper, we define them as persuasive techniques used by social engineers to gain access to conduct their social engineering attack. They are not attacks on their own, but a way to gain the target's trust to conduct a social engineering attack. Impersonation and pretexting are often used in conjunction with other social engineering attacks, such as email phishing [29,40], and is the psychosocial element used to trick targets, which is why we classify them as persuasive techniques, rather than social engineering attacks.

Impersonation:

is when the attacker pretends to be someone else to fool their target [8]. Impersonation is a persuasive technique because it persuades the target to trust the attacker. The target is being tricked into believing something that is not true, in order to make the social engineering attack successful.

In addition to the attacker pretending to be another individual, it can also refer to online items pretending to be something else to fool users. Users tend to judge sites on their "look and feel", which attackers can easily mimic [18, 65]. Unfortunately, users' mental models do not often align with phishing cues, such as browser icons, and sender names, which is why users are easily deceived [33].

Greed:

is often exploited, and is seen when attackers prey on users by offering rewards [5,33]. According to Kirlappos and Sasse, social engineering ultimately works because users want good deals [33]. Human weaknesses such as greed make users susceptible to falling for social engineering attacks [43].

The Need and Greed Principle, a term coined by Stajano and Wilson [54], states that users are susceptible to online scams because the scammers know what users want, and exploit these desires [54].

Fear:

is often used to exploit users by using threatening language, and threatening them with what may happen if they do not comply. Vulnerable populations and those who are scared of facing the repercussions of not complying are especially vulnerable to being exploited by fear tactics [5].

Scareware is an example of exploiting fear to get users to comply. It may claim a user's computer is infected with a virus, and then try persuading the user to purchase an unnecessary piece of software to remove it [28].

Curiosity:

is leveraged when an intriguing subject line or scenario is presented to potential targets [5]. The goal is that targets are intrigued enough to open the link/file, and be the victim of a social engineering attack. In a study on QRishing, it was shown that curiosity was the main motivator for users scanning unknown QR codes; they did not know what the QR code led to, and were curious to see what would happen [57].

3.3 Call to Action

The third stage of the framework is to include a call to action, or some sort of interaction where the target "initiates" the compromise. This means that there needs to be something tangible for the target to do which compromises them. Tangible actions include clicking a link, typing in a password, downloading a file, scanning a QR code, or giving up confidential information. Social engineering is effective because it simplifies the target's decision making; it nudges users towards the attacker's desired behaviours because figuring out what to do is straightforward [19].

This step was missing in the original framework. It is an important step because it is when users are beginning to comply with the social engineering attack. Once users

proceed to this step, they may potentially compromised their own, or their company's, security system by handing over confidential information, such as passwords or credit card information, or by installing malware.

3.4 Utilize the Information

The fourth step in this framework is when attackers use the information they have gained. It is one step towards achieving their target goal. In this stage, users are not interacting with anything. Instead, the attackers are using the information users gave for their own ends.

According to a 2020 Anti-Phishing Working Group (APWG) report, 86% of breaches were financially-motivated, and 37% targeted user credentials [2]. Obtaining company information, unauthorized access, conducting a social engineering attack as personal revenge, and collecting banking information are other common goals [2].

Depending on the type of resources the attack is targeting, and the scale of the attack, the attacker may only use the target's information once, or access it many times. For example, if the attacker targets credit card information, they may use the victim's credit card information many times to purchase items online until they are stopped. On the other hand, the attacker may target passwords for a website, and sell them, therefore they only use the information once.

3.5 Expands Scale

Social engineering attacks are spread to other individuals, or repeated with the same target(s). Some attacks, such as spear phishing because of its targeted nature, will require more follow-up than others, which means they will need to repeat back to gather more information.

The previous framework implied that the cycle only repeats with the same target. They did not indicate that the cycle could repeat with other targets, which needs to be addressed because most social engineering attacks can impact several individuals [11].

Chapter 4

User Mental Models of Social Engineering Attacks

4.1 Description

In this first study, we wanted to understand user mental models of social engineering attacks in relation to our proposed framework. We wanted to see where end users' mental models differed from the reality of social engineering, so we interviewed participants about their understanding of social engineering attacks.

We recruited 30 participants and held semi-structured interviews over video calls. We collected qualitative data from participants regarding their understanding of, and experiences with, social engineering attacks. We chose a qualitative approach for this first study because to our knowledge, there are no previous studies looking at mental models of the entire social engineering attack cycle, therefore we first need to understand what users understand about the topic, and a qualitative approach is appropriate for this.

The goal of our study is to understand gaps in users' understanding of social engineering attacks and compared them to using an updated attack framework. We seek to understand user mental models of the whole attack cycle, rather than one aspect of it, giving researchers a broader understanding of end user mental models.

4.2 Methodology

This study was conducted remotely using video calls with participants. First, participants were asked to fill out a demographics questionnaire (see Appendix A.1.1). The demographics questionnaire asked for gender identity, age, education, and self-rated knowledge of computer security. Additionally, it asked for participants' self-ratings of their susceptibility to social engineering attacks, which operating system they are

most familiar with, and how often participants used technologies such as social media, SMS text messaging, downloading mobile apps, and email. These questions were asked so that we could understand whether the operating system and participants' usage of technology impacts their experiences with social engineering.

After demographics were collected, we conducted semi-structured user interviews. Our questions were based on the steps of the framework. We asked participants questions about their understanding and experiences of social engineering. See Appendix A.1.2 for the interview questions. To avoid priming and leading participants, we were careful with the structure and order of the interview questions. We kept our questions open-ended, and avoided providing examples so participants' understandings of the question would not be influenced.

Participants were compensated $10 for a 30-minute interview. This study was cleared by Carleton University's Research Ethics Board (Clearance #115226), see Appendix A.1.3.

4.3 Participants

We recruited 30 participants using Prolific.co, an online crowdsourcing platform for recruiting research participants. Participants had to be over 18 years old, speak English, and not have a computer security background. Participants were from Canada and the US. The average age of participants was 28 years old ($SD = 8.47$). Our participants were relatively young, and well-educated. Most claimed to have some knowledge of computer security ($n = 27$), and most said they were somewhat susceptible to falling for phishing attacks ($n = 25$). Table 4.1 shows detailed participant demographics.

All participants reported using email daily, and the majority of participants reported using social media and SMS daily. Few participants reported downloading apps more than a few times a month.

Table 4.1: Participant demographics

		Count	Percentage
Gender	Female	17	56.7
	Male	12	40.0
	Non-binary	1	3.3
Current (or highest) level of education	High school	4	13.3
	Community college	3	10.0
	Bachelor's Degree	11	36.7
	Graduate or professional degree	11	36.7
	Trades	1	3.3
Self-reported computer security knowledge	Not at all knowledgeable	2	6.7
	Slightly knowledgeable	16	53.3
	Moderately knowledgeable	11	36.7
	Very knowledgeable	1	3.3
Self-reported phishing susceptibility	Very susceptible	3	10.0
	Moderately susceptible	12	40.0
	Slightly susceptible	13	43.3
	Not susceptible at all	2	6.7

4.4 Analysis

We used Thematic Analysis (TA) to analyze our qualitative results. Thematic analysis is a data analysis methodology which is used to systematically identify patterns within a qualitative data set [17]. We transcribed all audio recordings using Trint, and analyzed our data using NVivo 12. The lead researcher conducted the open coding process, looking for underlying patterns and ideas within the dataset. We organized our themes based on the stages of the framework. We also looked for any themes and codes which emerged outside of our model, and if there was anything that disproved it. The process of forming themes was done manually using sticky notes (see Figure 4.1).

4.5 Results: Open Coding

First, the interviews went through an open coding process, which was conducted by the lead researcher. In the open coding stage, we identified key ideas within our data,

and labelled them into codes. The researcher identified 59 codes, which fit into 12 themes. We were looking for patterns within our dataset that were related to the stages of our framework, and were open to findings outside of the framework. We understand that this is not the only way of organizing our data, and that the reason we organized our codes mainly around the stages of our framework is because they were intended to guide the questions and responses for our survey in the second study to further test our understanding of mental models of the social engineering attack framework. See Table 4.2 for a list of the open codes and their descriptions.

Table 4.2: The 59 codes used in the open coding process. The codes are organized into related groups based on the stages of the framework, in addition to other findings during the theme-finding process.

Code Name	**Description**
Stage 1a: Attack Vector	
Advertisements	Has seen ads being used for social engineering through pop-ups, side banner ads, or on social media.
SMS	Has seen SMS messages as a social engineering attack vector.
Email	Has seen emails as a social engineering attack vector.
Phone call	Has been called by a suspicious number, or someone (or a robocall) pretending to be someone else.
Websites	Has seen or been on a website they suspected to be a phishing website.
Social Media	Have seen or received social engineering through social media, often through ads or direct messages on Facebook, Instagram, Discord, Reddit, etc.
Stage 1b: Target Audience	
Anyone is susceptible	Anyone is susceptible to receiving social engineering attacks, and possibly falling for them. One social engineering attack is sent to thousands or more individuals.
Young teens	Group is more vulnerable to falling for social engineering than other groups, and attacks are adapting to this demographic.
Most vulnerable	Not everyone will fall for SE; some groups are more vulnerable.
Elderly	Elderly are more vulnerable to falling for social engineering attacks.

Continued on next page

Table 4.2 – *Continued from previous page*

Code Name	Description
Organizations	Companies and organizations can be targeted.
Why Users Received a Social Engineering Item	
Survey	Attackers use demographic surveys to gain personal information to contact targets.
Internet presence	Higher presence on the internet increases one's susceptibility to receiving social engineering attacks and for attackers to gain a better user profile to target attacks.
Tracking	Tracking one's internet and social media usage lets attackers send users social engineering attacks targeted to their preferences through ads, reward offerings, etc.
Stage 2: Persuasive Techniques	
Impersonation	Attackers impersonate another individual or organization.
Pretexting	Giving the target a fake story to get them to comply.
Curiosity	Attacks try to intrigue targets or cause confusion to get targets to comply.
Fear	Attackers use fear, such as threatening the target, or creating a sense of urgency.
Greed	Attacks offer something enticing to the target to get them to comply (e.g., offering money, or a free prize).
Social Engineering Attack Cues	
Intuition	User guided by intuition telling them if something is legitimate or not.
Call to action	Items asking users to click on a link or download something seem suspicious.
Too good to be true	If an offer is too enticing, it must be a scam.
Professionalism	An item's presentation, scenario, and whether the situation fits the user's situation indicates an item's legitimacy.
Spam filter	Users rely on a spam filter to filter out suspicious items.
Legitimate sender	Investigate the URL or phone number of the sender.
Post-hoc Social Engineering Attack Cues	
Ghosting	Once a user complies with the attacker's request, the attacker no longer responds to the user, or give the user their reward.

Continued on next page

Table 4.2 – *Continued from previous page*

Code Name	**Description**
Undo	It is difficult to undo the outcomes of a social engineering attack.
Notification	Notifications (or lack of) from third parties notifies users if it is a social engineering attack after they fell for it.
Stage 3: Call to Action	
Delete	User deletes suspicious items.
Ignore	User ignores suspicious items, avoids engaging with them.
Verify	User verifies to check if the item is legitimate (e.g., call the company, check the sender, Google it).
Click a link	User clicks the link presented.
User reports item	User reports suspicious items to the IT department, or the organization the attacker is impersonating.
Change password	User will change their account password upon receiving a suspicious item.
Mess with attackers	User pretends to fall for the attack to waste the attacker's time.
Share with others	If they believe the item is legitimate and offers a reward, users may share it with others.
Block sender	User blocks the sender to avoid being contacted again.
Attack Method	
Data breach	Data breaches to a company exposes a user to receiving social engineering attacks.
Malware	User downloads malware unknowingly by opening certain websites or email attachments.
Virus	Viruses as an indicator that one fell for a social engineering attack.
Automation	Use of bots or algorithms to run social engineering attacks quickly and send to as many targets as possible.
Reasons Users May (Not) Fall for Social Engineering	
Free time	More likely to engage with social engineering if user is bored.
Suspicion of everything	Anything online can be a social engineering attack, so users are careful with everything online.
Embarrassment	Users may not ask for social engineering help if they are embarrassed about having fallen for social engineering.
Others received it too	More likely to believe it is real if others received it too.

Continued on next page

Table 4.2 – *Continued from previous page*

Code Name	**Description**
Lack awareness	Users who lack awareness are more prone to falling for SE.
Awareness	Users keep up to date with things happening in technology or their life events.
Stage 4: Utilize the Information	
Personal Information	social engineering attacks collect personal information such as credit card, banking, password information.
Financial gain	Attackers try gaining money from targets, or sell their personal information for profit.
Hack	Attackers steal login credentials to hack into a target's accounts.
Attackers	
Attacker motivation	Attackers are driven by greed, desperation, or wanting power over targets.
Attacker identity	Attackers are from foreign (often developing) countries.
Attacker groups	Attackers may work solo, or in groups similar to underground crime organizations
Media portrayals	Users' views of attackers are influenced by media portrayals of hackers.
Stage 5: Expands Scale	
Forwarding	If an item seems legitimate, users may forward it to others to claim the promise of rewards.
Contacts	Attacks spread to one's contact list once a user falls for an attack.
Target many	Attacks target many users
Target specific	Attacks are specifically targeted to one user.
Networks	Attacks spread by spreading through networks (e.g., wifi network, all of one's accounts on their device, etc.).

4.6 Results: Themes

In this step, we drew connections between our open codes to identify underlying themes within our data. To understand end user mental models of social engineering attacks, we examined our open codes and realized that user *confidence* and *accuracy* were underlying themes within our data set. For different stages of the framework, users had different levels of confidence in their beliefs, and their beliefs had different levels of accuracy. In general, we found that users display high levels of confidence,

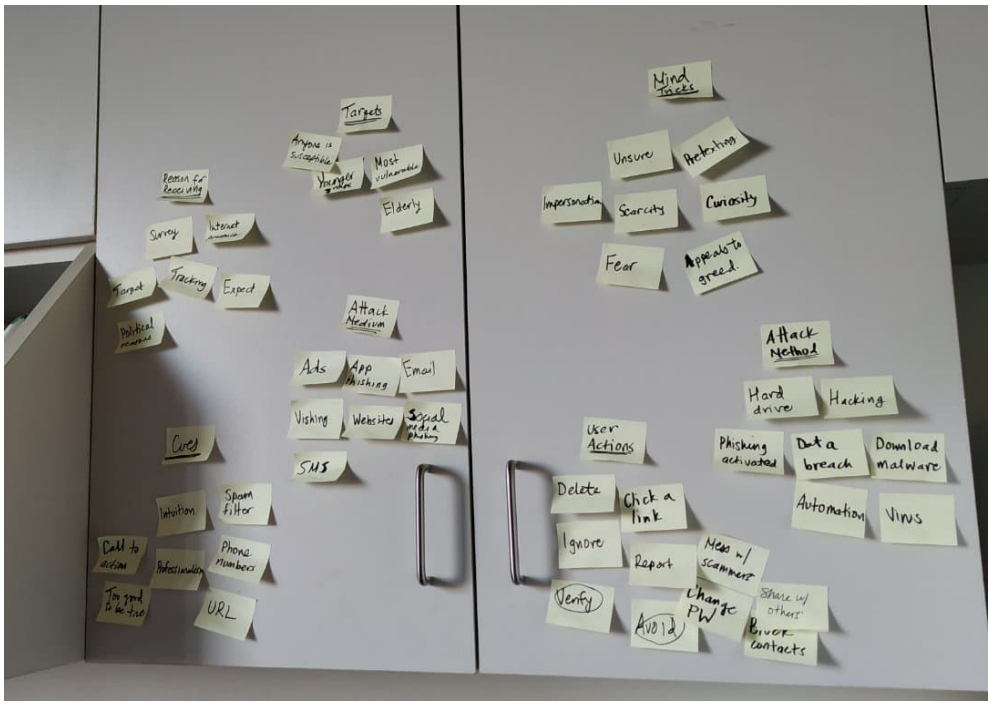


Figure 4.1: The manual theme-building process.

and lower rates of accuracy.

We organized our description of the themes around the stages of our proposed framework, and explained them according to the underlying themes of *confidence* and *accuracy*.

4.6.1 Stage 1: Attack Formulation

Attack Vector

Attack vectors refer to where users expect to encounter social engineering attacks. Participants expressed high confidence about being able to identify common social engineering attack vectors. Participants often mentioned specific attack vectors, such as emails and phone calls as common methods, but also frequently mentioned advertisements, SMS, websites, and social media as other attack vectors. These are attack vectors which participants are more careful with, and were suspicious of items they received.

Most participants associated social engineering with email, phone call, or SMS

phishing.

> "I always considered it to be more of an email thing... but recently I've been getting a lot of calls and text." – P15

We also had many participants list social media as another common attack vector.

> "(I expect to see it on) Facebook more than others, then email." – P9

However, participants showed that they also lacked some awareness about social engineering attack vectors. There were several emerging attack vectors which participants did not consider, such as app phishing, and not all participants were aware of the same kinds of social engineering attack vectors. Additionally, participants often showed incorrect understandings of certain attack vectors. For example, several participants had trouble differentiating between legitimate advertisements found on websites and social media, and phishing advertisements. Despite this, participants expressed confidence in their understandings about attack vectors.

> "There's a shopping feature on Instagram. So because of that, I think people will encounter (phishing) ads that pertain to what they were just browsing on Instagram or Facebook." – P16

This quote illustrates participants' over-confidence when describing their incorrect understandings of social engineering attack vectors. This participant was confusing targeted ads with phishing ads. We observed that some participants are inherently suspicious of internet tracking and see it as being invasive, but confuse attackers trying to breach security systems with companies trying to sell things.

For attack vectors, participants displayed high confidence in their beliefs, but their level of accuracy was not as high. This can result in participants having incorrect understandings of social engineering attack vectors, and may impact their behaviours when engaging with potential social engineering items.

Target Audience

This section of the *Attack Formulation* stage is about what users' perceptions about who is generally targeted in social engineering attacks, and why. In this section of the interview, participants expressed high confidence in their understanding about who gets targeted. Every participant said they had received a social engineering

attack in some form before. As a result, they often expressed beliefs that everyone is susceptible to receiving social engineering attacks, and that these attacks are sent to many individuals at once.

> "Everybody that is on the Internet or mobile phone (receives social engineering attacks)." – P9

Most participants were very confident that they themselves would not fall for a social engineering attack, citing their intuition or common sense to guide them.

> "A lot of the fraudulent ones (emails) seem obvious, you know, you I think it comes down a lot to common sense." – P30

Additionally, participants believed that very few users actually fall for social engineering attacks, which is correct [4,50]. This is because participants think most social engineering are very generalized attacks which are easy to identify, and are more of an annoyance.

> "I think they probably send out a mass email to specific people with accounts that they have on the list. And look what happens after they send out the mass email; I assume a very small percentage respond." – P21

Participants suspected that younger teens and the elderly are more vulnerable to actually falling for social engineering attacks. This belief is supported by the literature, which shows that the elderly and younger teenagers are more vulnerable to phishing [50].

Participants said the elderly are more vulnerable because they lack experience and knowledge about the Internet.

> "I think older people (are more vulnerable), because they don't necessarily understand how the Internet works and, you know, things that may be obvious to us, they obviously have been working on." – P1

On the other hand, participants mentioned that younger individuals are also vulnerable because they lack life experience, and are naive.

> "It's most likely to be really young people because they don't have any experience in using payments on the Internet or actually going through the process of checking out for the purchase." – P3

Participants also said that the only time users who are not part of particularly vulnerable groups may fall for social engineering attacks is when the phishing scenario fits a user's situation, and causes them to comply due to fear, or other persuasive techniques.

> "I saw one recently that told me my stimulus check was lost and I needed to provide my Social Security Number, which obviously I would never receive in an email. I had already gotten the stimulus check, but if someone's wondering where their check is, they might put in their information, their social security number, their name, thinking that it's a legitimate email." – P24

Some participants mentioned a major target for social engineering attacks: larger organizations, such as businesses, hospitals, and universities. Participants were mainly focused on end users being the targets of social engineering attacks, but some mentioned that organizations may also be subject to social engineering attacks. Users believed organizations could be targets because attackers would gain access to a lot more personal information, and the potential financial gains are much greater if they target an organization rather than individuals.

> "In the context of a health care setting, a thief could be after an employee's network credentials to be able to know if I work for a hospital, for example, and log into a network that allows them access to the computer network and gives them resources. A thief might want to obtain those resources to be able to implant malware on the network so that they can hold the resources of the network hostage to be able to gain Bitcoin from them." – P28

Why Users Received a Social Engineering Item Participants described that attackers had a list of contact emails, names, and/or phone numbers to contact their targets with, and that if attackers have access to this information, they will send social engineering attacks to them. Participants were confident that attackers got ahold of user contact information from surveys, tracking users, and through monitoring a user's internet presence.

Participants believed malicious surveys which ask for more information than necessary, and those asking for private information such as credit card details, could be

used by attackers to gain personal information. Surveys could be either used to gain contact information to further contact the user, and/or used to collect more personal information for attackers to do something more malicious with.

> "Maybe social media, like 'Click here to sign up' or fill out some kind of petition or something to that collects information about you." – P17

Participants also showed their suspicion of targeted social media ads, and believed they are linked to surveys collecting personal information about them to send social engineering attacks.

> "You see the ad on social media and it appeals to you, and you click on it and provide your information like your email or phone number, I think that's how they'll (the attackers) reach you through those means as well." – P16

Participants showed confidence that their internet presence, which refers to a user's level of activity on the internet, as another reason for why they keep receiving social engineering items. Participants believed that the more "present" they were online, the more likely attackers could collect personal information from them to contact them, or to entice them with a targeted social engineering attack. As a result, participants believed that if one were to use the internet less often, they would be safer from receiving social engineering items.

> "I mean, it has to do with tracking our Internet activities. So I might be more active than most. And then therefore, they'll catch my signature and follow up with me while I'm surfing (the internet)." – P9

Due to this belief that higher internet presence exposes one to potential social engineering attacks, and general suspicion of how user data is being used, participants commonly mentioned trying to be careful with what they engage with online, and trying not to put their information out there for companies or attackers to use.

> "I use social media. But as a reader, I don't post a lot. I typically use it as a way to get news, sort of an advance information tool. I'm not so much, you know, a person that posts on social media. So I'm not as vulnerable as other people might be." – P28

Similar to internet presence, participants were confident that another reason they received social engineering items was because of online user tracking. Participants

believed that attackers could be tracking their internet usage, and as a result, could be sent social engineering items masquerading as targeted ads, messages/emails, or reward offerings to entice them.

> "Maybe from Googling, cookies can give info to people about your habits and maybe there are things that you are interested in at that specific time and they can offer you things at really low prices that will trick you into actually buying." – P3

P3 shows an inaccurate understanding of the difference between legitimate targeted ads and phishing ads which reflects the views of other participants. They do not trust internet user tracking in general, and view all kinds of tracking and targeted messages as suspicious.

4.6.2 Stage 2: Persuasive Techniques

Participants displayed high confidence in their ability to identify common persuasive techniques used by attackers. Participants most commonly believed that attackers will try appealing to greed most often by offering something enticing to the target, such as money, a free reward, or information the target may be interested in.

> "By email, sometimes it's a call where there's something really cool article or product, and you're like, 'Wow, I want to look at this more' and then it tries to lead you somewhere else." – P9

This participant believed that most email phishing attacks try deceiving users by appealing to their curiosity and/or greed. They will provide an enticing offer, and lead users elsewhere once they fall for the bait.

Participants also believed that attackers may also provoke a sense of fear in users, often by pretending to be a person or organization with authority, such as the government or police, so that users feel they must comply.

> "They develop a sense of urgency, so much like the CRA (Canada Revenue Agency) is coming for you, you better call me back now and give me $10, 000, or call up the FBI squad, so people get scared." – P10

Attackers may make users curious by providing a scenario or message that is appealing to users. When users are curious, they are more likely to engage with the social engineering attack item to further investigate it, according to participants.

> "I saw an ad on YouTube, and clicked on it. The ad was offering nutrition advice and tips, but the website wasn't related to that." – P20

Many participants mentioned that attackers will use impersonation to trick users into complying. Through impersonation, they can create a false presence that fools the user into believing they are someone else.

> "So for phishing, you're talking about creating fake profiles, using bots, images, and trying to create a false sense of an individual, trying to create a sense of, I guess, a feeling for an individual and also trying to create a friendship, a partnership, a relationship, some sort of connection." – P1

Participants also said that attackers can give them a pretexting scenario which also tricks them into complying with the request.

> "So they come up with a story, the one we all know about is the Nigerian prince and stuff like that. But there are more elaborate stories now. It's like, 'Your car will be impounded and we need your Social Security.' They come up with a story that puts you in immediate danger so that you have to respond to the email. So, yeah, I think that the story is the way that they really get at you." – P7

Participants understood that different persuasive techniques target different vulnerabilities, and when used in the right context can be persuasive at getting users to comply.

Social Engineering Attack Cues

Participants were confident in their ability to detect social engineering attacks. Out of our 30 participants, only two participants admitted outright they have trouble with determining whether an item is legitimate or not (P4, P12). Participants commonly mentioned relying on their intuition, spam filter, or looking at cues within the social engineering item, such as its level of professionalism, whether the sender is legitimate, whether the social engineering item has a call to action, and if an offer is too good to be true as cues that an item is malicious.

Participants who were confident in their ability to detect social engineering items relied on their intuition to tell them if something was legitimate or not. Due to the high prevalence of phishing [11], and phishing-related warnings users receive, users may have ingrained ideas of what phishing looks like, and are therefore automatically

able to classify these items. However, this intuition mainly applies to email, social media, and phone call phishing attacks, which are common attack vectors [2, 11, 14].

> "I look at them and it's like second nature. I don't really give them piece of mind too much." – P1

Users also said they rely on technical mitigation strategies, such as spam filters, to help them detect social engineering attacks. Email, SMS, and sometimes social media often have a spam filter, which users rely on to filter out most social engineering attacks. For the majority of attacks, users said their spam filter did a good job of identifying and filtering these out.

> "Most of the time, my email accounts automatically will sort spam out, but I've gotten a few in the last few days that I've had to manually mark as spam." – P24

Sometimes, social engineering attacks show up in their main inbox, and that is where they rely on other cues to tell if it is a social engineering attack of not. Professionalism, which refers to a social engineering item's look and feel, spelling and grammar, legitimacy of the message, and whether the request makes sense, was commonly cited by participants as a cue for social engineering.

Participants may look out for professionalism in the message content, such as whether the call to action makes sense.

> "If they asked me to pay them in a really obscure away like 'Pay me $100, but it has to be through Vanilla MasterCard cards, Bitcoin or some kind of untraceable way' is weird. Netflix isn't going to ask you to pay using Bitcoin or on Pizza Pizza gift cards." – P10

Professionalism was also commonly mentioned in terms of the look and feel of an item, such as the spelling and grammar, or graphics used looking professional.

> "So as you get an email, it claims to be something and it's like the grammar is not right or there's something spelled wrong is usually a pretty clear indicator that it's not a legitimate company." – P15

Participants also mentioned checking the sender of the suspicious item to verify whether it is a social engineering attack or not. They described checking the email address, company name, or cell phone number to see whether these line up with what is legitimate.

> "I would double check the sender's email address because at first it'll say Scotiabank, but when you do click on it, it's like a bunch of random numbers and letters." – P11

When a suspicious item includes a call to action, such as encouraging recipients to click a link, download a file, or send payment, participants took this as an additional cue that an item is malicious. Some users said they will click these suspicious links to further investigate their validity.

> "Sometimes there's also the link they send you, if you go click on the link, you'll see that it's not from the official Scotiabank or Facebook or Instagram." – P11

Some participants also reported not opening any links through email to prevent any kind of harm, even if it is from a legitimate entity. Instead, they will Google what the link claimed to be and log in through the known website, or enter in the website URL in their browser.

> "If it's something (email) from my bank or that appears to be from my bank asking me to do something, you know, I'll ignore the links in the email and go directly to the bank site to do it." – P30

Lastly, participants believed that if an offer is too attractive, it probably is a malicious item. Many participants (or their friends and family) expressed complying with, or almost complying with a request because the offer was attractive, but most participants claimed to be suspicious of any offer that was too appealing.

> "I've seen on social media like Facebook, there are fake job postings. Sometimes they post in public university groups. It just doesn't seem legitimate because if you're from Canada, and there are these people from the US looking for Canadian students to work part time, and the pay is extremely high" – P5

For some participants, this suspicion extends to legitimate businesses offering rewards to customers; some participants believed this is similar to a social engineering attack, and avoid complying with getting these rewards.

> "A lot of stores will have reward systems. I know Wal-Mart and Harris Teeter had a few to fill out a certain survey, then they enter you into a grocery sweepstakes or something. And sometimes the links might take you to like a separate place that has nothing to do with it. And they just want you to fill out demographic information." – P14

Post-hoc Social Engineering Attack Cues

In addition to social engineering Attack Cues that users may assess as they receive suspicious items, this section describes social engineering Attack Cues that users consider after potentially engaging with the item. Participants displayed the lowest confidence when trying to explain their beliefs about how to tell if they fell for a social engineering attack.

Participants said that a major clue is that the attacker will "ghost" the target. Once the target complies with the attacker's request, they will stop responding to messages, and the reward will never arrive.

> "I was continuously being excited just because this person was responding, and they sent a picture of the tickets and they were keeping up with it. And then afterwards (after I sent money), they stopped responding, and I became confused and the panic started to settle in." – P4

Participants also note that another clue is that the impacts of a social engineering attack are difficult to undo. Once a user sends personal information, money, or installs malware, it is difficult to get one's information or money back, or remove the malware.

> "You provide them the information, and then once you've done that, it's very, very difficult to undo it because it's there. You know, they're in an unknown location and they're using a variety of mechanisms to cloak their their location, et cetera." – P28

Lastly, participants also believed that notifications signify that they fell for a social engineering attack. There will either be a lack of notifications from a legitimate source verifying that a user's actions were recognized, especially in the case of password changes, or a notification from a legitimate source warning the user of strange activity, such as banks indicating there were strange purchases made on one's credit card.

> "Suddenly there are charges on your credit card or your credit score suddenly takes a dive when you haven't changed anything yourself." – P24

Participants also noted that they rely on their anti-virus software to alert them of threats; if their software finds a virus, they will assume they fell for a social engineering attack.

> "I try to update all my security settings and I run a scan, but I just assume they're in there." – P9

4.6.3 Stage 3: Call to Action

During this stage, participants expressed confidence about how social engineering attacks are initiated, but were not as accurate in their beliefs. Participants had mixed understanding about when exactly they themselves became vulnerable to a social engineering attack.

Some participants said they have been compromised as soon as they are sent a social engineering attack item.

> "If I get an email or something, I think, 'OK, where did I go wrong along the line before this to lead to me getting this in the first place.'" – P15

Others believed it is when one opens a social engineering attack item.

> "When I open the email, it downloads something to my computer that reads my email or searches for the information they're looking for." – P17

Most said it is when one engages with a social engineering attack, such as clicking links, downloading a file, or giving out the requested personal information.

> "I reckon if they sent me a link and I clicked on that link, that's when bad things start to happen. They can either yoink some kind of information from me or I'll get put on a list saying, 'OK, this guy is actually gullible enough to open the email so we should send him more emails.'" – P10

If participants suspect an item is a phishing attempt, their actions upon receiving the item differ from those who believe it is legitimate.

When participants believe an item is fraudulent, they will often report it to the IT department (if it exists), the platform which was used to carry out the attack, or notify the person/organization the attacker is impersonating.

> "I've reported things on Instagram, like on social media. If you get direct messages from bots and things like that, I'll report those just because that's a very accessible and easy way to potentially handle it." – P15

Some participants also mentioned deleting these attack items. P27 describes both deleting these items, and changing their password upon receiving social engineering items.

> "I delete it from my email. I delete from the trash and I change my password." – P27

Many participants said they ignore suspicious items, and avoid opening and engaging with them.

> "I don't open any emails from people I don't recognize." – P17

Additionally, participants may also block the sender to avoid getting contacted by them again. This is often the case for suspicious phone calls, SMS texts, and emails.

> "If I keep repeatedly getting a call from a certain number of text messages, I'll block the number to prevent myself from getting it." – P15

Sometimes, participants also say they may mess with the attackers for fun. They will pretend they are unaware they are engaging with a social engineering attempt, and try wasting the attacker's time.

> "Usually I try and mess with them for as long as I can. I've been getting these DHL scams on my phone recently. I try to mess with them, waste their time. You know, if I'm driving, I love to do that because I can multitask and have some fun at the same time." – P2

When participants are unsure about the legitimacy of an item, they also attempted to verify the item, such as by checking the URL and sender, Googling it, or asking family and friends, and sometimes clicking the link presented in the item.

> "I usually just try to look up the email address of the sender or if it's a phone number, I try to look at the phone number. And if it matches with if it matches with anything real, then I figure, OK, this is legit. But if it doesn't, then I just block and move on." – P23

When users believe the social engineering item's claims, they report complying with its call to action, and also sharing it with others if it offers an enticing reward.

> "They'll send an email from our local pizzeria. They'll be like 'Send this to five people and you get free pizza for a year' or something like that, and I've done that. It turned out it was a scam. My mom actually told me not to send it anymore because once you click that link, it wasn't anything about the free pizza. And I didn't know. I just was trying to be helpful." – P12

In this stage of the framework, participants understood that they generally should not click any unfamiliar links or download malware, but they lack awareness of exactly when a social engineering attack is initiated. This subsequently impacts their behaviours when acting upon the social engineering item, and impacts how they familiarize themselves with accurately identifying social engineering attacks.

Reasons Users May (Not) Fall for SE

In this section, users generally said that others may fall for social engineering attacks, but they themselves were not likely to for them. In these questions, we observed that most participants excluded themselves from risk; they mentioned how others may be prone to falling for SE, but mentioned how they protect themselves from these attacks.

Users typically believed that those who fall for social engineering attacks often do because of four possible reasons. Firstly, the user may have had free time; when users have free time and are bored, they may be more likely to explore the social engineering item, and engage with it. Secondly, users may be embarrassed, either because they do not want to ask someone if an item is legitimate or not, or because they fell for a social engineering attack and do not want to ask for help. Thirdly, if others around the user received a social engineering item, the user may be more likely to trust it. Lastly, a user may lack awareness of social engineering attacks, and may lack awareness of how to protect themselves.

> "If you don't spend time on the Internet or if you don't really know how the government works and things like that you would have known that the CRA (Canada Revenue Agency) doesn't really doesn't call you or that Netflix doesn't do things like this." – P10

On the other hand, there are two main reasons users mentioned as reasons for why they do not fall for social engineering attacks. First, users say they try to be cautious and suspicious of most things they encounter on the Internet. These users claim to verify suspicious items they receive, are hesitant to give out personal information online, and try to avoid clicking unfamiliar links.

> "I'm pretty hesitant about what I interact with online because I do feel like a lot of things online, it has a higher tendency to be something like that (phishing)." – P8

Secondly, users also commonly mentioned that they prevent themselves from falling for social engineering attacks by keeping up to date with their life situation, such as keeping an eye on bank records, or events in their family or life so they are aware of their situation.

> "I keep good eye on my credit so that I know there's been nothing that has hit against my my personal credit." – P30

Attack Method

Participants did not express as much confidence or accuracy in their understandings of social engineering attack methods. Many were unsure of exactly what happens after they initiate an attack; if no money was lost, they were unsure about what happens to them or their device(s) if they fall for an attack.

> "There would be no way to tell (if I fell for an attack) because if they're not asking me for money, then I don't know what the point of the phishing attack is either." – P10

As such, there were different beliefs, such as if social engineering attacks are attempted through data breaches, or by installing malware. Participants held incorrect beliefs, or got social engineering mixed up with other security attacks.

> "If you're shopping for something and you provide an email and password and credit card information for that website, that website may be subject to a phishing attack and so on." –P13

For example, P13 associated phishing only with data breaches that happen to organizations. They did not know that they could compromise themselves, and thought that it was only when a data breach occurred that their information could get compromised.

4.6.4 Stage 4: Utilize the Information

This stage of the framework is not visible to users, and users do not interact with anything here, therefore what happens during this stage is mainly based on participants' speculations about what attackers do with their information. Participants were confident in their understanding that the goal of social engineering attacks is to gain personal information of users, hack into user accounts, and/or for financial gain. Participants believed that regardless of whether attackers take personal information, or hack into accounts, their ultimate goal is to turn this into financial gain for themselves.

> "I guess I always kind of think that the information they're going for is financial. I think either they would use it directly or they would sell the information to someone else." – P17

There is a belief that attackers will target personal information such as credit cards, bank information, or Social Security/Insurance Numbers. Many participants also frequently mentioned the dark web with regard to personal information. They believed attackers might try to sell personal information on the dark web, or add their information to a database of other collected information to share with other attackers.

> "(Attackers are) looking to access your accounts directly or posting this information on to the dark web or something for other individuals to access it." – P25

These participant beliefs are accurate, as previous research has shown that social engineering attacks commonly target financial resources and/or personal information, and sometimes hack into accounts [14, 34].

Attackers

Participants held stereotypical beliefs about the attackers running social engineering attacks. Participants often believed they were motivated by greed, desperation, or a sense of power over targets, and that attackers worked either on their own, or in large groups similar to underground criminal organizations.

Participants were confident that attackers were generally from foreign (often developing) countries. Participants believed attackers were from developing countries because running social engineering attacks can be very lucrative, and provide a source of income for them.

> "You think about the Chinese, North Korea, Russia. These are countries that sure, they may be well off, but are they really? In terms of a lot of the freedoms that they (the citizens) have, and the opportunities they have to provide for their families." – P1

Participants believe this because the social engineering attacks often contained many spelling and grammar mistakes, and phone calls with speakers with foreign accents, so they assumed attackers must not be native English speakers in their view.

> "(I don't picture) necessarily any specific ethnicity or race, but foreign because sometimes you get calls and the voices are clearly foreign." – P15

Participants mentioned that their ideas of attackers are often influenced by media portrayals of hackers. As a result, their idea of what hackers are trying to accomplish, where they come from, what they target impacts how they view attackers running social engineering attacks.

> "I would say without being too stereotypical, based on what I've read in the news and so forth, (attackers) typically come from foreign countries. So typically, like Russia, China, some from the Middle East." – P25

4.6.5 Stage 5: Expands Scale

In this stage, participants expressed high confidence in their beliefs about attack scalability. Participants commonly believed social engineering attacks are spread when users forward emails to others, when the attack spreads to one's contacts, or through computer networks, such as wifi networks or to one's other accounts.

> "You get one computer infected, it remotely infects other computers on the network. And once you get that computing power, it's just like a robotic army." – P1

Participants commonly said that social engineering attacks are generalized, and target many individuals at once. They believed thousands or more individuals are targeted in each attack because very few fall for these attacks, so it is important to cast a wide net.

> "I would suspect that they are sent to many people and they are trying to maybe raise the percentage of people that might actually go through the process and give them what they want." – P3

At the same time, participants also have a conflicting belief that some social engineering attacks, especially advertisements, are specifically targeted to an individual user. Participants who have this belief state that companies or attackers are tracking their online activities, and try to entice them into falling for a social engineering attack by presenting something they might be interested in:

> "Facebook and Instagram are together (owned by the same company), so if you're logged in and you have certain information about you that's public, I feel like in some way, a lot of (phishing) websites have some sort of access to that so you'll see certain things on both social media platforms like the phishing ads which will either be on Facebook or Instagram or both because the two are connected." – P16

Despite participants displaying high confidence in their beliefs, they do not have an accurate understanding of how social engineering attacks scale. They have inconsistent beliefs about how many users these attacks target, and their reasoning behind spearphishing attacks is confused with general internet user tracking.

4.7 Mental Models of Social Engineering Attacks

The results of this first study suggest that end user mental models of social engineering attacks are impacted by user confidence and the high prevalence of social engineering attacks. Every participant said they had seen social engineering attacks, and that they see them very frequently. As a result of this high prevalence of being exposed to social engineering attacks, users often think they understand them, and have high confidence in their understanding.

Most users do not get extensive formal training on phishing or social engineering awareness, so most mental models are formed by previous experience. The high prevalence of social engineering attacks users encounter impacts what users associate with SE, such as how it looks like, what to do when one encounters an attack, and what it targets. Unfortunately, we have also observed from this study that participants also confuse social engineering with a variety of other concepts, such as targeted advertisements, other security attacks, spam, and fraud. Participants have threat-level misunderstandings, and often believe social engineering is less dangerous, and more of a nuisance, rather than an actual threat to companies and systems. Not all threats are at the same level, but participants do not seem to understand the different levels in threats, which may be why they are so confident in their beliefs. In essence, users may believe social engineering is less threatening, and more of a nuisance, hence their confidence in their abilities to deal with social engineering.

Since most participants believed they experienced high degrees of success with detecting and avoiding social engineering attacks, this feeds into their high confidence levels about their understanding of these attacks.

4.7.1 Confidence and Accuracy

We observed that participants generally displayed high levels of confidence in their understanding of social engineering attacks. Despite this high confidence, their level of accuracy was much lower. Some stages of the framework, such as *Stage 1: Attack Framework* and *Stage 2: Persuasive Techniques* had higher confidence and accuracy, but other stages such as *Stage 5: Expands Scale* had much lower levels of accuracy, but consistently high confidence amongst participants.

Most notably, participants had misunderstandings about what social engineering is. Participants confused social engineering with fraud, spam, and other security attacks. This misunderstanding can possibly lead users into wrongly evaluating the threat levels of social engineering attacks, and believing they are a nuisance, rather than an actual threat to their security.

This ties in with participants' beliefs that individuals, rather than organizations, are often the targets of social engineering attacks. Their beliefs about how they were initially targeted with a social engineering item conflicts with their beliefs about the scale of social engineering attacks, as mentioned in Stage 5's results. Participants generally believed social engineering attacks are generalized attacks targeting as many users as possible, and tend to target individual users, rather than organizations. Yet, it is evident that they also think social engineering attacks are highly targeted to users based on their personal information and internet browsing. Participants displayed confidence in their beliefs about why they may have received a social engineering item, yet also held inconsistent, and often inaccurate beliefs about social engineering attacks.

Chapter 5

Examining Users' Confidence and Accuracy by Attack Stage

5.1 Description

In the first study, we found that participants were confident in their abilities to detect and understand social engineering attacks, but held some important misunderstandings about them. To further investigate, we conducted a second study to evaluate and quantify participants' confidence and accuracy in their understanding of social engineering attacks to understand where in the attack framework users have the most difficulty. The first study looked at understanding broader user mental models of social engineering, and this study looked at quantifying how the themes of confidence and accuracy are evaluated using a larger sample of participants. This second study quantifies participants' confidence judgements and accuracy within different stages of the social engineering attack framework, and seeks to more explicitly identify where users are struggling when it comes to understanding social engineering attacks.

5.2 Methodology

In the second study, we used a survey to quantitatively measure how confidence and accuracy play a role in mental models of social engineering attacks. We investigated how confident and accurate participants are for the different stages of our proposed framework, to understand where users tend to have difficulty when it comes to social engineering attacks.

In our survey, participants were presented with screenshots of real or fabricated social engineering items, and then answered questions about these attacks. The questions and their multiple-choice responses were based off of the stages of the framework, and the themes and codes developed from the first study.

We sourced a total of 21 different attack examples from attacks encountered by

the researcher, and family and friends. We also sourced additional examples from online blogs and industry social engineering reports which showed examples of social engineering items received by others.

We aimed to have social engineering attack examples representing a wide variety of scenarios, and different levels of professionalism and realism to portray the wide variety of social engineering attacks that exist. These examples also represented the wide variety of social engineering attacks mentioned in our framework; we had examples of targeted and generalized attacks, examples which showed the various kinds of persuasive techniques, different calls to action, and different audiences. We included a variety of attacks so that participants could see a diverse sample representing a wide variety of attacks.

We collected roughly equal numbers of email, SMS, advertisement, and social media social engineering examples. The attack vectors, meaning the medium in which the attack is normally encountered, selected were influenced by the first study; we chose attack vectors participants said they have seen so we could properly represent what users were seeing in their own life. A.1 in the Appendix describes the 21 examples we showed participants. Figures 5.1 and 5.2 explain some of the examples we included in the survey, and explains how they fit into our social engineering framework.

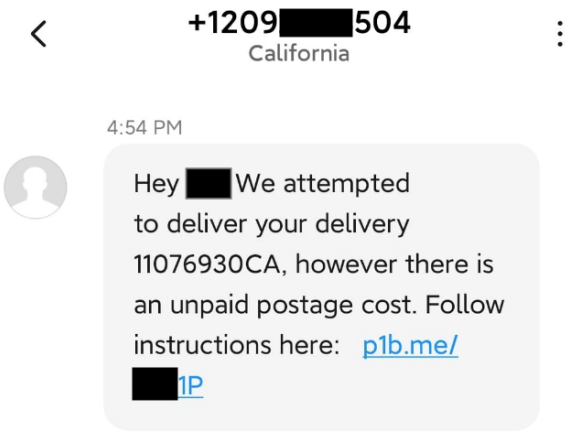


Figure 5.1: Screenshot of an SMS social engineering attack we showed participants. This example is a generalized attack trying to elicit fear to gain money from users.

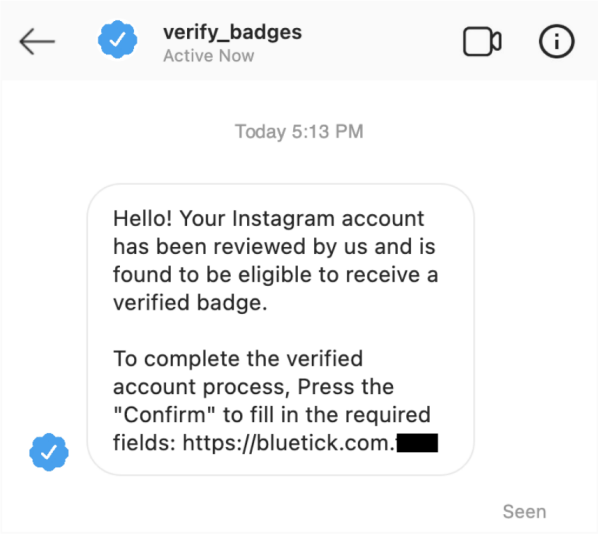


Figure 5.2: Screenshot of a social media social engineering attack we showed participants. This example is a targeted attack trying to appeal to greed to gain money.

To prevent participant fatigue, we randomly presented 7 of the 21 examples to each of our 150 participants. Since the order in which examples were shown was randomized and no one saw all 21 examples, each example was viewed and evaluated by 52 participants. Participants were clearly told these were attacks, so they were instructed to not interact with the actual attacks, nor to enter any personal information for these attack evaluations. Participants were compensated $3 for 20 minutes. The study was cleared by Carleton University's Research Ethics Board (Clearance #115762), see Appendix A.2.3.

Based on these attack examples, participants were asked questions such as who they think they attack is being sent to, whether the attack is generalized or targeted, whether they believe others would fall for it, and ratings of their confidence that their answer was accurate. We had one question per example representing each stage of the attack framework. For each example, we asked five accuracy questions testing how participants understood how the stages of the framework applied to the example. The questions and choices were the same, but the answers differed based on the specific example. For instance, the question corresponding to Stage 1 (Attack Formulation: Target Audience), the question was "Who is this item most likely being sent to?", and the options were: "Anyone", "Specific group based on shared interests, demographics,

etc.", and "Just the one recipient." Following each question, participants were asked to rate how confident they were about their response, ranging from 1 (Not Confident) to 5 (Very Confident). In addition to questions relating to our framework, we also had multiple choice questions about questions outside of our framework, such as "How susceptible do you think you are to falling for this attack?" and "How susceptible do you think others are to falling for this attack?".

5.3 Participants

We used Prolific.co, a research participant recruitment website, to recruit 150 participants for this survey. Two failed our attention check, so we replaced them with two other participants.

Table 5.1 shows a detailed description of the participants. We recruited 69 women, 78 men, and 3 non-binary participants. They were relatively young; over half (111/150) were under 35. Most participants (117/150) claimed to have "some" computer security knowledge. Most participants' highest level of completed education was high school (56/150), followed by a Bachelor's degree (44/150), then graduate or professional school (37/150). Participants reported that they used social media most frequently (132/150 used it daily), then SMS (124/150), and then email (108/150).

As for phishing vulnerability, most participants believed their likelihood of falling for phishing was low; only 18/150 said they were extremely or very likely to fall for it, and 26/150 said they were not likely at all. Most participants have not received any phishing training (121/150), but of those who did, most did the training out of their own interest (19/150), or due to workplace obligations (10/150). Twenty-three participants reported they previously have been victims of phishing, 81/150 said they have not been victims, and 46/150 were unsure.

5.4 Hypotheses

Based on the results from our first study, we formed three hypotheses for this second study.

H1: *Self-reported confidence rates will differ based on the stages of the framework.*

Table 5.1: Participant demographics

		Count	Percentage
Gender	Female	69	46.00
	Male	78	52.00
	Non-binary	3	2.00
Current (or highest) level of education	High school	56	37.33
	Community college	10	6.67
	Bachelor's Degree	44	29.33
	Graduate or professional degree	37	24.67
	Trades	2	1.33
	No school completed	1	0.77
Age	18 - 25	90	60.00
	26 - 35	21	14.00
	36 - 45	30	20.00
	46 - 55	8	5.33
	55+	1	0.77
Self-reported computer security knowledge	Not at all knowledgeable	19	12.67
	Slightly knowledgeable	62	41.33
	Moderately knowledgeable	55	36.67
	Very knowledgeable	11	7.33
	Extremely knowledgeable	3	2.00
Self-reported phishing susceptibility	Extremely susceptible	2	1.33
	Very susceptible	16	10.67
	Moderately susceptible	36	24.00
	Slightly susceptible	70	46.67
	Not susceptible at all	26	17.33

In our first study, we observed that participants displayed varying levels of confidence for different aspects of social engineering. For example, we noticed that participants were less confident when describing what happens once they comply with a social engineering attack, but highly confident when identifying the various social engineering attack vectors. Therefore, although participants generally displayed high confidence in the first study, we believe this second study will more precisely show that there will be variations in confidence through the framework, and where confidence is different.

After each question measuring accuracy, there was a Likert-scale question asking

participants to evaluate how confident they were that their response was correct. Participants rated their confidence from 1 (Not Confident) to 5 (Very Confident).

We examined confidence scores by the stages of the framework, rather than by example because we wanted to compare confidence across the entire framework.

H2: *Accuracy rates of social engineering attack evaluations will differ based on the stages of the framework.*

Based on the results of the first study, we noticed that participants consistently held some beliefs that were inaccurate, and also some which were accurate. Participants often got social engineering confused with online fraud, spam, and online advertisements. However, beliefs about persuasive techniques used on social engineering seemed largely correct, and participants could identify the different ways in which attackers may manipulate users.

To determine the correct answers for each social engineering attack example for the survey, we relied on industry security reports which described how these attacks worked, and also had two researchers assess the examples. In the survey, participants filled in multiple-choice questions to give their evaluations of different social engineering attacks. These questions corresponded to a different stage of our proposed framework, and each had a different number of possible question options. For example, the accuracy question for Stage 1 was "Who is this item most likely being sent to?" and the options for ever example were "Anyone", "Specific group based on shared interests, demographics, etc.", and "Just the one recipient." Every accuracy question had the same choices, but the correct example differed based on the example being shown.

Following each question, there was a question asking participants to evaluate how confident they were in their response.

We aggregated participant responses to each other across all the social engineering examples, and compared the correct responses to the participants' responses to calculate how accurate they were across the study. Once again, we did not look at accuracy by social engineering example, but rather across the stages of our framework.

H3: *There will be interaction effects between confidence and accuracy based on the stages of the framework.*

From our observations in the first study, we hypothesize that when confidence and accuracy are combined, there will be interaction effects which will appear only when combined.

We aggregated the confidence and accuracy scores for each set of questions in the survey that represented a stage of the framework by combining participants' scores for each accuracy question relating the framework and its corresponding confidence question.

Participants were assigned an aggregate score from 1 to 10; a score closer to 1 means that they rated themselves as highly confident in their response, but were incorrect, and a score closer to 10 meant participants rated themselves as high confident in their response and were correct. For example, a participant selected "Very confident" but got their answer wrong would get a score of 1, and a participant who got their answer right but felt "Moderately confident" would get a score of 7.

5.5 Results

5.5.1 Confidence

In general, confidence scores among participants were quite high. For each of the confidence questions corresponding to a stage in our proposed framework, the mean confidence score was 4, "Confident" (see Table 5.2). To see if there were any significant differences in the confidence ratings between the different stages of the framework, we ran a Kruskal-Wallis test since our data was non-parametric. We found that there were no significant differences between the different stages ($H(4) = 3.06, p = 0.55$).

Table 5.2: Confidence results

Stage of the Framework	**Median**	***SD***
S1: Attack Formulation (Target Audience)	4	1.01
S2: Persuasive Techniques	4	0.96
S3: Call to Action	4	0.96
S4: Utilize the Information	4	1.03
S5: Expands Scale	4	0.97

Confidence ratings: 1 = Not confident, 5 = Very confident

H1: *Self-reported confidence rates will differ based on the stages of the framework.*

We did not find support for this hypothesis, as the confidence ratings were the same across all stages of the social engineering framework. Confidence remained high, which indicates that users tend to be more confident in their abilities to avoid and detect social engineering attacks than they actually can.

5.5.2 Accuracy

We found that overall, accuracy rates were quite low for all stages of the framework. Table 5.3 shows the accuracy rates based on the different stages of the framework, and Figure 5.3 illustrates these differences in accuracy rates.

Table 5.3: Accuracy results

Stage of the Framework	**Count (out of 1050)**	**Percentage Correct**
S1: Attack Formulation (Target Audience)	367	35%
S2: Persuasive Techniques	406	39%
S3: Call to Action	207	20%
S4: Utilize the Information	362	34%
S5: Expands Scale	485	46%

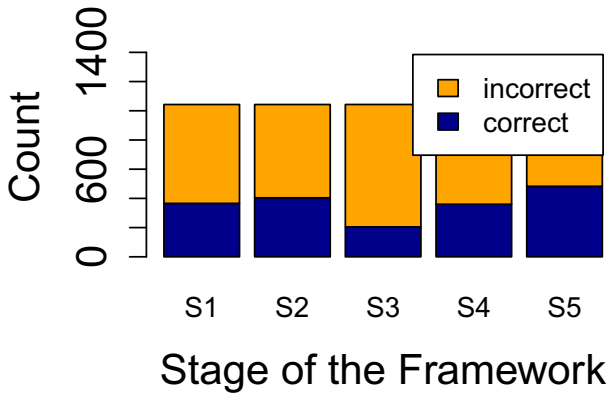


Figure 5.3: Accuracy rates by the stages of the framework.

Table 5.4: Probability of accuracy: Chi-squared results

Stage of the framework	**Correct**	**Number of question options**	X^2	p
S1: Attack Forumlation (Target Audience)	35%	3	0.001	1
S2: Persuasive˜ Techniques	39%	4	0.001	1
S3: Call to Action	20%	5	0.001	1
S4: Utilize the Information	34%	3	0.001	1
S5: Expands Scale	46%	2	0.001	1

Since we used a multiple choice survey for the accuracy scores, one possibility is that participants may be guessing in their responses. We observed that participants' accuracy rates were often related to the probability of an answer being correct. We ran Chi-squared tests with a Bonferroni correction and a 0.05 alpha level to compare how different the participants' accuracy scores were from the probability of being correct to determine if participants were guessing. None of the accuracy scores were significant, which indicates that there was no evidence that participants performed better than if they were guessing at any stage. Table 5.4 displays the accuracy scores compared to the chances of any possible answer being correct, and the Chi-squared test results.

H2: *Accuracy rates of social engineering attack evaluations will differ based on the stages of the framework.*

We did not find support for this hypothesis because participants did not perform significantly better than if they were guessing. Despite participants' guessing, they expressed high confidence in their possible answers, revealing that participants overrate their confidence in their abilities to understand social engineering attacks.

5.5.3 Confidence and Accuracy Interactions

To examine the interaction between confidence and accuracy, we aggregated our confidence and accuracy scores, and compared how they fared based on the different stages in our proposed framework. To calculate the aggregate score, the accuracy score was coded as either 1 (if correct) or -1 (if incorrect), and was multiplied by the confidence

score, to give a value between either 1 to 5, or -1 to -5. We wanted a coherent ordinal scale, so we shifted everything upward and mapped the 1–5 scores to 6–10, and the -1 to -5 scores to 5 to 1, which reversed the previously-negative numbers.

Scores from 1 to 5 on the x-axis are incorrect responses, and scores from 6 to 10 are correct. A response close to 1 means participants rated themselves as highly confident in their response, but were incorrect, and a response close to 10 meant participants rated themselves as high confident in their response and were correct. Since our data did not meet assumptions of normality (see Figure 5.4), we conducted non-parametric statistical tests.

To determine whether significant differences existed between the different aggregated scores based on the stages of the framework and attack vector, we ran a Kruskal-Wallis test. We found that there were significant differences between the stages of the framework ($H(4) = 129.25, p < 0.001$).

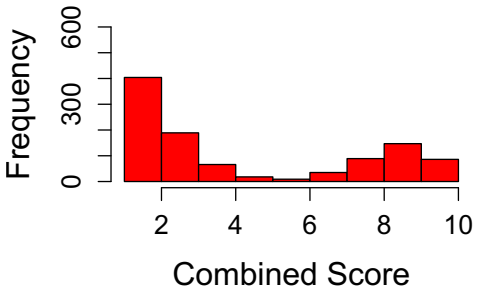
To see where these differences were, we ran pairwise Wilcoxon rank sum tests with a Bonferroni adjustment to correct for multiple comparisons. Table 5.5 shows the results for the Wilcoxon tests at the different stages of the framework.

Table 5.5: Wilcoxon results by the stages of the framework

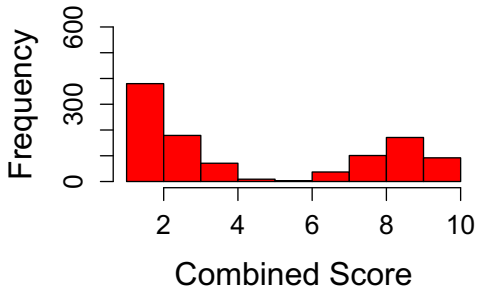
	S1: Attack Formulation	**S2: Persuasive Techniques**	**S3: Call to Action**	**S4: Utilize the Information**
S2: Persuasive Techniques	$U = 520942$ $p = 0.09$	–	–	–
S3: Call to Action	$U = 633532$ $p < 0.001$**	$U = 656452$ $p < 0.001$**	–	–
S4: Utilize the Information	$U = 536884$ $p = 0.60$	$U = 560284$ $p = 0.23$	$U = 444660$ $p < 0.001$**	–
S5: Expands Scale	$U = 487976$ $p < 0.001$**	$U = 511426$ $p = 0.02$*	$U = 398196$ $p < 0.001$**	$U = 493388$ $p < 0.001$**

* indicates significance at the $p < 0.05$ level, ** indicates significance at the $p < 0.001$ level.

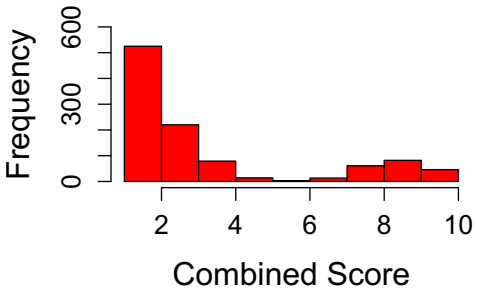
Overall, our results showed that there were not many differences in confidence and accuracy, except for Stage 3 (*Call to Action*) and Stage 5 (*Expands Scale*). Participants performed significantly worse for Stage 3, and significantly better for Stage 5 compared to the other stages.



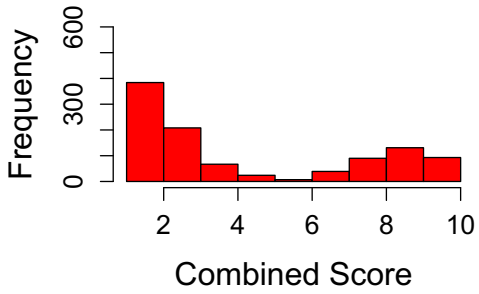
(a) S1: Attack Formulation



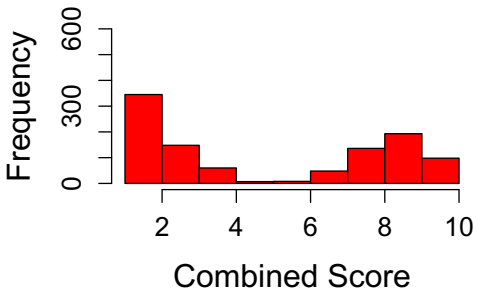
(b) S2: Persuasive Techniques



(c) S3: Call to Action



(d) S4: Utilize the Information



(e) S5: Expands Scale

Figure 5.4: Distributions of aggregated confidence and accuracy scores by the stages of the framework.

We believe that the discrepancies for Stages 3 and 5 may be due to guessing, as we explained in the *Accuracy* section. Since the question for Stage 3 had 5 options, and stage 5 had 2 options, the likelihood of getting a correct answer for Stage 5 was much higher.

Alternatively, Stage 3 may have lower scores because participants had knowledge about this particular stage. Phishing research has shown that users tend to not be very knowledgeable about phishing, and what to do when they encounter phishing attacks [18, 59, 68]. There are many ways to get phished, and it is not always clear when a user exactly falls for a phishing attack. For some attacks, the attack is initiated when they click a link, but for others, receiving a pop-up ad indicates that their computer is already infected, and other attacks may only be initiated once the participant goes through a pays the fee. Social engineering attacks have a deliberate lack of transparency, which makes it difficult for end users to know what to do when they encounter different attacks, especially when it comes to unfamiliar attacks [20].
H3: *There will be interaction effects between confidence and accuracy based on the stages of the framework.*

We found support for this hypothesis, as confidence and accuracy significantly differed at different stages, specifically between Stage 3 and Stage 5.

5.5.4 Exploratory Results

To investigate analyses outside of our *a priori* hypotheses, we ran some exploratory analyses to investigate if any findings emerged outside of our original hypotheses.

Demographics Findings

We ran Kruskal-Wallis tests to find if there were any significant differences for confidence and accuracy scores between different demographic groups. We compared the aggregated confidence and accuracy scores by gender, whether participants received phishing training previously, and whether they were a victim of phishing. We found that there were no significant differences by gender ($H(2) = 0.72, p = 0.70$), previous phishing training ($H(1) = 0.42, p = 0.52$), or if they were a previous phishing victim ($H(2) = 3.71, p = 0.16$).

Since all of these results were non-significant, we did not investigate any further.

Attack Vector Findings

We ran a Kruskal-Wallis test, comparing the aggregate confidence and accuracy scores by attack vector, and found that there were significant differences ($H(3) = 278.07, p < 0.001$). Since the result of this test was significant, we ran Wilcoxon rank sum tests comparing the aggregated scores by attack vector. See Table 5.6 for detailed results by attack vector.

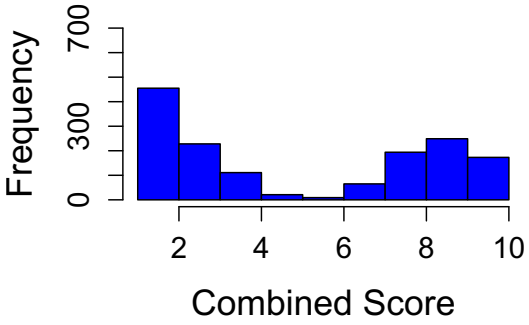
Table 5.6: Wilcoxon Results by Attack Vector

	SMS	**Email**	**Advertisement**
Email	$U = 1092788$ p <0.001**	–	–
Advertisement	$U = 985342$ $p = 0.03$*	$U = 683576$ p <0.001**	–
Social Media	$U = 1239902$ p <0.001**	$U = 879379$ p <0.001**	$U = 979682$ p <0.001**

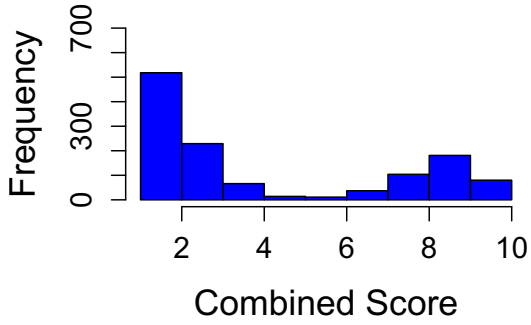
* indicates significance at the p <0.05 level, ** indicates significance at the p <0.001 level.

Compared to other attack vectors, participants overall had significantly lower scores for social media attacks. Despite the popularity of email social engineering attacks, we were surprised to find that email had significantly lower aggregated scores compared to SMS and advertisement-based social engineering attacks. Participants performed the best when it came to SMS, then advertisement attacks, then email, and performed the worst for social media attacks. Figure 5.5 visualizes the scores by attack vector.

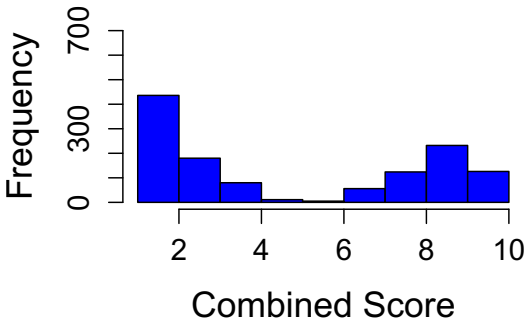
We believe participants performed especially poorly on social media attacks because this is an emerging attack vector, therefore not all participants will be aware of how it looks like or how it works. For one, social media phishing allows for highly personalized attacks to occur [58], which may be unexpected for participants, therefore they may be unsure of how to respond. Additionally, social media is constantly evolving at a very fast rate, which makes it harder for users to quickly become experts in [58, 64].



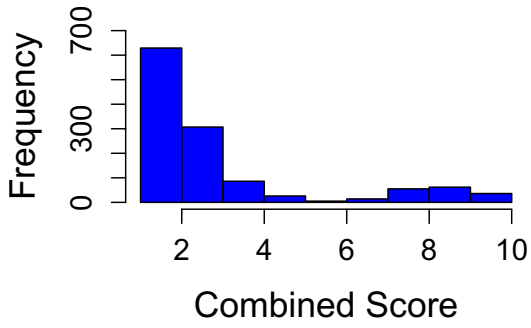
(a) SMS



(b) Email



(c) Advertisements



(d) Social Media

Figure 5.5: Distributions of aggregated confidence and accuracy scores by attack vector.

Chapter 6

Discussion

Social engineering is one of the most common IT security attacks [4] because users are easier to manipulate than breaking into a security system [41, 61]. To prevent users from falling for social engineering attacks, several technical and educational mitigation strategies have been implemented [22, 23, 36–38, 51]. Despite the high prevalence of social engineering, and the variety of mitigation strategies which exist, it still remains a problem which many users still fall victim to [2, 3, 50].

The research problem this thesis investigated was: what are end users' mental models of social engineering attacks? We believe that to create user-centric social engineering mitigation strategies, we must first understand user mental models of social engineering attacks. To investigate this problem, we proposed an updated social engineering attack framework, building on Mitnick's [41] and Mouton et al.'s [42] previous social engineering frameworks. We used our social engineering framework as the basis for the two studies and to investigate how users' mental models align with how social engineering attacks function. Users are overly confident in their ability to understand and detect social engineering attacks. This has implications for users' mental models, as it leads them to believe their mental models of social engineering are generally accurate.

We conducted two studies. In the first study, we interviewed 30 participants to ask them about their experiences and perceptions of social engineering. Using thematic analysis, found that the underlying themes of the first study were that participants seemed to have high confidence in their beliefs, yet also displayed some inaccurate beliefs.

To further investigate how *confidence* and *accuracy* impact mental models of social engineering, we ran a second survey with 150 participants. Participants looked at various social engineering attack examples, and filled in a survey asking for their

understanding of these attacks based on the different stages of the framework. Participants' evaluations were scored for accuracy and confidence. The results showed that confidence levels remained high throughout the different stages, whereas accuracy was low throughout, and that participants were most likely guessing.

Many users were confident in their abilities to accurately detect and understand how social engineering attacks work due to the high prevalence of social engineering attacks they encounter. Having high confidence in inaccurate beliefs is dangerous because it prevents users from seeking new information about social engineering attacks, and leads them to continue using outdated mental models to guide their responses to social engineering. This is especially dangerous with regards to emerging social engineering attack vectors where many users said they did not know existed as potential attack vectors.

Participants' accuracy scores in the second study were related to the probability of getting an answer right by guessing. However, despite the probability of guessing, participants were still confident in their answers. This reveals that users tend to be overconfident in their sometimes inaccurate beliefs about social engineering.

Almost every participant in the first study was confident they could detect phishing attempts, and said they had never fallen for one. However, despite this confidence, several participants mentioned clicking links, and filling in surveys to further investigate suspicious items, indicating that they were unknowingly complying with social engineering attacks. Participants excluded themselves from risk, and often mentioned other groups they are not part of, such as the elderly, or young teens, as being at risk. In the second study, we observed a similar pattern; only 18/150 said they were very or extremely susceptible to falling for social engineering attacks.

Participants' overconfidence can be explained by the Dunning-Kruger effect, which states that those who are less knowledgeable about a concept tend to overestimate their actual ability [35]. In both studies, we did not have any participants with a computer science or computer security background, and not many self-reported as being very knowledgeable about computer security. Despite this, participants believed they were unlikely to fall for phishing attacks.

This mismatch between users' ideas of their own vulnerability compared to that

of others may be an example of the fundamental attribution error, which states that individuals attribute their own successes to their own abilities, and others' successes to their environment and surroundings [48]. The majority of users were unaware of the impacts of social engineering attacks, so participants might have thought their ability to avoid and detect social engineering attacks was due to them understanding these attacks well, and knowing how to avoid them. However, for other users, participants might have thought these users were able to avoid these attacks due to luck, or some external factors.

In our first study, we observed that participants were not always sure about what constituted social engineering. Most participants tended to only think of phishing as a social engineering. Some participants confused social engineering with online fraud, other security attacks, spam, and online advertising. These participants would describe legitimate practices, such as targeted ads and entering surveys to win prizes, as phishing methods, and also show confidence in their beliefs. This awareness and caution of companies selling personal data for profit extends into user mental models of social engineering. Users may get these confused, and sometimes cannot distinguish between when companies are collecting user information to sell things, or when attackers are collecting info for illegal and malicious reasons. Users' beliefs about social engineering are influenced by data collection and advertising practices. Many phishing techniques can also be used in legitimate (though dubious) contexts as well. For example, multi-level marketing schemes often appeal to greed, offer appealing things to sellers, and may have account names for representatives which are different from the main account. Legitimate advertisements also try appealing to greed and fear, offer things which sound too good to be true, and use personal data to target ads towards users.

When it came to understanding the target audience of an attack, there was also confusion among participants for both studies. On one hand, participants commonly believed social engineering attacks were generalized, and sent to many individuals, regardless of their specific demographics and interests. However, participants would also contradict their statements by believing that social engineering attacks are also very targeted to just an individual. Certain attack vectors, such as advertisements

and social media, were more likely to be flagged as very targeted attacks by participants. Notably, participants tended to think that end users were the main targets of social engineering attacks, leaving out organizations and larger systems. As a result, participants had misconceptions regarding the threat levels of social engineering attacks, and tended to believe social engineering was more of an annoyance for users to deal with, rather than a threat to security systems.

Overall, our studies show that there are important gaps between user mental models and current social engineering mitigation strategies. These gaps in knowledge, paired with high confidence may result in resistance to learning more about phishing, which can leave users vulnerable.

6.1 Design and Education Recommendations

Our studies suggest a number of ways in which design and education opportunities which can be implemented based on our results. In both studies, participants displayed high confidence and low accuracy across the stages of a social engineering attack. Participants had misunderstandings about who and what social engineering attacks target, the scale and severity of these attacks, and the various social engineering attack vectors.

To address the mismatch between confidence and accuracy, social engineering education campaigns should focus more on explicitly telling users they may be more vulnerable than they think. Current education tends to focus on identifying phishing items [36, 51], but based on the results of our studies, we believe it should focus more on teaching users about what happens when they interact with a social engineering item, letting them know when in the attack cycle they are most vulnerable, and how to avoid spreading the item to others.

Social engineering education should also focus on clearing up misconceptions regarding the severity of these attacks. Participants generally believed end users were the main targets of social engineering attacks, and found social engineering to be less threatening compared to other security attacks because they view them as a nuisance consisting of ads, emails, etc. that most users can easily identify. If a user is also an employee of an organization, carrying this misconception can have negative impacts

on the organization. Therefore, users should be taught the different ways in which social engineering can be a threat.

We acknowledge that these educational propositions will be difficult to implement, especially given that users are prone to the fundamental attribution error when evaluating social engineering attacks. Users may not realize that they are also prone to poor habits, or realize how vulnerable they are. Therefore, emphasizing users' vulnerabilities to social engineering will need to be done in a way which is mindful of user biases.

Lastly, many participants in the first study incorrectly confused social engineering with spam, fraud, online advertising, and other security attacks. This is possibly due to users seeing similarities between social engineering and other concepts, such as how they may be a nuisance, trying to trick users, or value personal data, but the lines between what is legitimate and not are unclear to them. Therefore, we suggest that websites and mobile apps could include a signal for legitimate items. Similar to verified social media accounts, "verified" emails, ads, social media posts, etc. can act as a signifier to users that an item is legitimate, and not a phishing item. However, we also note that attackers can also spoof legitimate cues, and creating unspoofable legitimate cues can be difficult and may not be recognized by users. Previous research on phishing and security icons has shown that many users are unaware of these cues, or misunderstand them [18, 20, 23].

Chapter 7

Conclusion

Social engineering is a common security attack which targets users' psychological vulnerabilities [34, 41]. Unfortunately, current social engineering mitigation strategies do not fully prepare users for detecting and avoiding social engineering attacks.

In this thesis, we proposed an updated social engineering attack framework to better understand modern attacks, and conducted two studies to understand mental models of social engineering attacks using this framework as a guideline. The first study consisted of interviews with 30 participants to collect qualitative data about mental models of social engineering attacks, and the second study consisted of a survey with 150 participants who evaluated various social engineering items to quantify how confidence and accuracy play a role in mental models, and to investigate where users tend to struggle the most.

We found that participants have major misconceptions about social engineering which, when paired with their high confidence, can result in increased user vulnerability. We suggest that there may be ways for education and design to mitigate these problems, for example by verifying legitimate advertisements, and educating users about the severity of social engineering attacks.

7.1 Limitations

Due to the nature of Prolific, our participants skewed towards being younger and better educated. As a result, this group may be more technologically savvy, and more aware of phishing attacks compared to other demographic groups. However, despite the younger demographic of our studies, participants did not appear to be as successful at understanding social engineering attacks.

Since we did not want to influence our participants' understandings of social engineering, not all participants were clear about what phishing or social engineering

were, so it is possible we are missing some information about what they understand about the topic.

In both studies, we did not have participants interact with actual social engineering attacks to protect them from potential consequences. We did not investigate how participants would evaluate and behave with social engineering attacks in real life. Instead, we relied on self-reports about how participants would typically behave, which might not be representative of reality and can sometimes be unreliable. Additionally, there is a lack of context which participants were given with the social engineering cues, since they only viewed screenshots. This lack of context may have impacted participants' accuracy scores for the second study as well.

7.2 Future Work

Our studies showed that participants tended to underestimate the threat levels of social engineering attacks. In the future, it would be interesting to see further research on user perceptions of social engineering attack threat levels. Additionally, we found that participants confused social engineering with online advertising, spam, fraud, etc. A study further exploring what users count as social engineering, and why they believe it is social engineering may provide further insights into end users' mental models.

Bibliography

[1] About Let's Encrypt - Let's Encrypt - Free SSL/TLS Certificates.

[2] APWG | Phishing Activity Trends Reports.

[3] Google Safe Browsing – Google Transparency Report.

[4] RSA Quarterly Fraud Report: Q1 2019.

[5] Sherly Abraham and InduShobha Chengalur-Smith. An overview of social engineering malware: Trends, tactics, and implications. *Technology in Society*, 32(3):183–196, 2010.

[6] Luca Allodi, Tzouliano Chotza, Ekaterina Panina, and Nicola Zannone. The need for new antiphishing measures against spear-phishing attacks. *IEEE Security & Privacy*, 18(2):23–34, 2019.

[7] Chaitrali Amrutkar, Patrick Traynor, and Paul C Van Oorschot. Measuring SSL indicators on mobile browsers: Extended life, or end of the road? In *International Conference on Information Security*, pages 86–103. Springer, 2012.

[8] Scott D Applegate. Social engineering: Hacking the wetware! *Information Security Journal: A Global Perspective*, 18(1):40–46, 2009.

[9] Ken Bain. *What the best college students do*. Harvard University Press, 2012.

[10] Cristian Bravo-Lillo, Lorrie Faith Cranor, Julie Downs, and Saranga Komanduri. Bridging the gap in computer security warnings: A mental model approach. *IEEE Security & Privacy*, 9(2):18–26, 2010.

[11] Mark Button and Cassandra Cross. Technology and fraud: The 'fraudogenic' consequences of the internet revolution. *The Routledge Handbook of Technology, Crime and Justice. London: Routledge*, 2017.

[12] Cristiano Castelfranchi and Rino Falcone. Trust is much more than subjective probability: Mental components and sources of trust. In *Proceedings of the 33rd Annual Hawaii International Conference on System Sciences*, pages 10–pp. IEEE, 2000.

[13] Pew Research Center. Mobile fact sheet. *Internet & Technology*, 2017.

[14] RSA Anti-Fraud Command Center. RSA quarterly online fraud report, 2020.

[15] Rui Chen, Joana Gaia, and H Raghav Rao. An examination of the effect of recent phishing encounters on phishing susceptibility. *Decision Support Systems*, 2020.

[16] Robert B. Cialdini. *Influence: The Psychology of Persuasion*, volume 55. Collins New York, 2007.

[17] Victoria Clarke and Virginia Braun. Thematic analysis. In *Encyclopedia of Critical Psychology*, pages 1947–1952. Springer, 2014.

[18] Rachna Dhamija, J Doug Tygar, and Marti Hearst. Why phishing works. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 581–590, 2006.

[19] Xun Dong, John A Clark, and Jeremy Jacob. Modelling user-phishing interaction. In *2008 Conference on Human System Interactions*, pages 627–632. IEEE, 2008.

[20] Julie S Downs, Mandy B Holbrook, and Lorrie Faith Cranor. Decision strategies and susceptibility to phishing. In *Proceedings of the Second Symposium on Usable Privacy and Security*, pages 79–90, 2006.

[21] Maeve Duggan and Joanna Brenner. *The demographics of social media users, 2012*, volume 14. Pew Research Center's Internet & American Life Project Washington, DC, 2013.

[22] Serge Egelman, Lorrie Faith Cranor, and Jason Hong. You've been warned: an empirical study of the effectiveness of web browser phishing warnings. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1065–1074, 2008.

[23] Adrienne Porter Felt, Robert W Reeder, Alex Ainslie, Helen Harris, Max Walker, Christopher Thompson, Mustafa Embre Acer, Elisabeth Morant, and Sunny Consolvo. Rethinking connection security indicators. In *Twelfth Symposium on Usable Privacy and Security*, pages 1–14, 2016.

[24] Diksha Goel and Ankit Kumar Jain. Mobile phishing attacks and defence mechanisms: State of art and open research challenges. *Computers & Security*, 73:519–544, 2018.

[25] Slade E Griffin and Casey C Rackley. Vishing. In *Proceedings of the 5th Annual Conference on Information Security Curriculum Development*, pages 33–35, 2008.

[26] Christopher Hadnagy. *Social Engineering: The Art of Human Hacking*. John Wiley & Sons, 2010.

[27] Tzipora Halevi, Nasir Memon, and Oded Nov. Spear-phishing in the wild: A real-world study of personality, phishing self-efficacy and vulnerability to spear-phishing attacks. *Phishing Self-Efficacy and Vulnerability to Spear-Phishing Attacks*, 2015.

[28] Ryan Heartfield and George Loukas. A taxonomy of attacks and a survey of defence mechanisms for semantic social engineering attacks. *ACM Computing Surveys (CSUR)*, 48(3):1–39, 2015.

[29] Tom N Jagatic, Nathaniel A Johnson, Markus Jakobsson, and Filippo Menczer. Social phishing. *Communications of the ACM*, 50(10):94–100, 2007.

[30] Anna Kang, Jae Dong Lee, Won Min Kang, Leonard Barolli, and Jong Hyuk Park. Security considerations for smart phone smishing attacks. In *Advances in Computer Science and its Applications*, pages 467–473. Springer, 2014.

[31] Athanasios Karakasiliotis, SM Furnell, and Maria Papadaki. An assessment of end-user vulnerability to phishing attacks. *Journal of Information Warfare*, 6(1):17–28, 2007.

[32] Christopher M Kelley, Kyung Wha Hong, Christopher B Mayhorn, and Emerson Murphy-Hill. Something smells phishy: Exploring definitions, consequences, and reactions to phishing. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 56, pages 2108–2112. SAGE Publications, Los Angeles, CA, 2012.

[33] Iacovos Kirlappos and M Angela Sasse. Security education against phishing: A modest proposal for a major rethink. *IEEE Security & Privacy*, 10(2):24–32, 2011.

[34] Katharina Krombholz, Heidelinde Hobel, Markus Huber, and Edgar Weippl. Advanced social engineering attacks. *Journal of Information Security and applications*, 22:113–122, 2015.

[35] Justin Kruger and David Dunning. Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of personality and social psychology*, 77(6):1121, 1999.

[36] Ponnurangam Kumaraguru, Justin Cranshaw, Alessandro Acquisti, Lorrie Cranor, Jason Hong, Mary Ann Blair, and Theodore Pham. School of phish: A real-world evaluation of anti-phishing training. In *Proceedings of the 5th Symposium on Usable Privacy and Security*, pages 1–12, 2009.

[37] Ponnurangam Kumaraguru, Yong Rhee, Steve Sheng, Sharique Hasan, Alessandro Acquisti, Lorrie Faith Cranor, and Jason Hong. Getting users to pay attention to anti-phishing education: Evaluation of retention and transfer. In *Proceedings of the Anti-Phishing Working Group's 2nd Annual eCrime Researchers Summit*, pages 70–81, 2007.

[38] Ponnurangam Kumaraguru, Steve Sheng, Alessandro Acquisti, Lorrie Faith Cranor, and Jason Hong. Teaching Johnny not to fall for phish. *ACM Transactions on Internet Technology (TOIT)*, 10(2):1–31, 2010.

[39] Xin Robert Luo, Stephen Burd, Wei Li, Richard G Brody, William B Brizzee, and Lewis Cano. Flying under the radar: Social engineering. *International Journal of Accounting & Information Management*, 2012.

[40] William R Marczak and Vern Paxson. Social engineering attacks on government opponents: Target perspectives. *Proceedings on Privacy Enhancing Technologies*, 2017(2):172–185, 2017.

[41] Kevin D Mitnick and William L Simon. *The Art of Deception: Controlling the Human Element of Security*. John Wiley & Sons, 2003.

[42] Francois Mouton, Mercia M Malan, Louise Leenen, and Hein S Venter. Social engineering attack framework. In *2014 Information Security for South Africa*, pages 1–9. IEEE, 2014.

[43] Gareth Norris, Alexandra Brookes, and David Dowell. The psychology of internet fraud victimisation: A systematic review. *Journal of Police and Criminal Psychology*, 34(3):231–245, 2019.

[44] Gunter Ollmann. The vishing guide. *IBM Technical Report*, 2007.

[45] Gregory L Orgill, Gordon W Romney, Michael G Bailey, and Paul M Orgill. The urgency for effective user privacy-education to counter social engineering attacks on secure computer systems. In *Proceedings of the 5th Conference on Information Technology Education*, pages 177–181, 2004.

[46] Jens Rasmussen and Aage Jensen. Mental procedures in real-life tasks: A case study of electronic trouble shooting. *Ergonomics*, 17(3):293–307, 1974.

[47] Jens Riegelsberger, M Angela Sasse, and John D McCarthy. The mechanics of trust: A framework for research and design. *International Journal of Human-Computer Studies*, 62(3):381–422, 2005.

[48] Lee Ross. The intuitive psychologist and his shortcomings: Distortions in the attribution process. In *Advances in experimental social psychology*, volume 10, pages 173–220. Elsevier, 1977.

[49] Anna L Rowe and Nancy J Cooke. Measuring mental models: Choosing the right tools for the job. *Human Resource Development Quarterly*, 6(3):243–255, 1995.

[50] Steve Sheng, Mandy Holbrook, Ponnurangam Kumaraguru, Lorrie Faith Cranor, and Julie Downs. 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*, pages 373–382, 2010.

[51] Steve Sheng, Bryant Magnien, Ponnurangam Kumaraguru, Alessandro Acquisti, Lorrie Faith Cranor, Jason Hong, and Elizabeth Nunge. Anti-phishing phil: the design and evaluation of a game that teaches people not to fall for phish. In *Proceedings of the 3rd Symposium on Usable Privacy and Security*, pages 88–99, 2007.

[52] Steve Sheng, Brad Wardman, Gary Warner, Lorrie Cranor, Jason Hong, and Chengshan Zhang. An empirical analysis of phishing blacklists. 2009.

[53] David Silver, Suman Jana, Dan Boneh, Eric Chen, and Collin Jackson. Password managers: Attacks and defenses. In *23rd USENIX Security Symposium*, pages 449–464, 2014.

[54] Frank Stajano and Paul Wilson. Understanding scam victims: Seven principles for systems security. *Communications of the ACM*, 54(3):70–75, 2011.

[55] Keith S Taber. Mediating mental models of metals: Acknowledging the priority of the learner's prior learning. *Science Education*, 87(5):732–758, 2003.

[56] Paul C van Oorschot. *Computer Security and the Internet*. Springer, 2020.

[57] Timothy Vidas, Emmanuel Owusu, Shuai Wang, Cheng Zeng, Lorrie Faith Cranor, and Nicolas Christin. QRishing: The susceptibility of smartphone users to QR code phishing attacks. In *International Conference on Financial Cryptography and Data Security*, pages 52–69. Springer, 2013.

[58] Arun Vishwanath. Getting phished on social media. *Decision Support Systems*, 103:70–81, 2017.

[59] Melanie Volkamer and Karen Renaud. Mental models–general introduction and review of their application to human-centred security. In *Number Theory and Cryptography*, pages 255–280. Springer, 2013.

[60] Rick Wash and Molly M Cooper. Who provides phishing training? Facts, stories, and people like me. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1–12, 2018.

[61] Rick Wash and Emilee Rader. Influencing mental models of security: a research agenda. In *Proceedings of the 2011 New Security Paradigms Workshop*, pages 57–66, 2011.

[62] Tara Whalen and Kori M Inkpen. Gathering evidence: Use of visual security cues in web browsers. In *Proceedings of Graphics Interface 2005*, pages 137–144. Canadian Human-Computer Communications Society, 2005.

[63] Daniel T Willingham. *Why don't students like school?: A cognitive scientist answers questions about how the mind works and what it means for the classroom*. John Wiley & Sons, 2009.

[64] Ryan T Wright and Kent Marett. The influence of experiential and dispositional factors in phishing: An empirical investigation of the deceived. *Journal of Management Information Systems*, 27(1):273–303, 2010.

[65] Min Wu, Robert C Miller, and Simson L Garfinkel. Do security toolbars actually prevent phishing attacks? In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 601–610, 2006.

[66] Kuldeep Yadav, Ponnurangam Kumaraguru, Atul Goyal, Ashish Gupta, and Vinayak Naik. SMSAssassin: Crowdsourcing driven mobile-based system for SMS spam filtering. In *Proceedings of the 12th Workshop on Mobile Computing Systems and Applications*, pages 1–6, 2011.

[67] Affan Yasin, Rubia Fatima, Lin Liu, Awaid Yasin, and Jianmin Wang. Contemplating social engineering studies and attack scenarios: A review study. *Security and Privacy*, 2(4):e73, 2019.

[68] Olga A Zielinska, Allaire K Welk, Christopher B Mayhorn, and Emerson Murphy-Hill. Exploring expert and novice mental models of phishing. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 59, pages 1132–1136. SAGE Publications Sage CA: Los Angeles, CA, 2015.

Appendix A

Appendix

A.1 User Study 1

A.1.1 Demographics Questionnaire

1. What is your gender identity?
    - Male
    - Female
    - Non-binary
    - Prefer not to say
2. How old are you?
3. How knowledgeable are you about computer security?
    - Not knowledgeable at all
    - Slightly knowledgeable
    - Moderately knowledgeable
    - Very knowledgeable
    - Extremely knowledgeable
4. What is your highest level of education?
    - High school
    - Community college
    - Trades
    - Undergraduate degree
    - Graduate or professional degree
5. What is your occupation?
6. How susceptible do you think you are to phishing (social engineering)?
    - Extremely susceptible
    - Very susceptible
    - Moderately susceptible

- Slightly susceptible
- Not susceptible at all

7. What kind of smartphone do you own?
    - iPhone
    - Android
    - Don't own a smartphone
    - Other: please specify
8. What kind of computer operating system are you familiar with?
    - macOS
    - Microsoft Windows
    - Both Windows and macOS
    - Don't own a laptop
    - Other: please specify
9. How often do you use email?
    - Daily
    - Sometimes (a few times a week)
    - Rarely (less than a few times a month)
    - Never/don't have email
10. How often do you use social media?
    - Daily
    - Sometimes (a few times a week)
    - Rarely (less than a few times a month)
    - Never/don't have social media
11. How often do you download new apps?
    - Daily
    - Sometimes (a few times a week)
    - Rarely (less than a few times a month)
    - Never/don't have a device to download apps to
12. How often do you use SMS (text messaging) on your phone?
    - Daily
    - Sometimes (a few times a week)
    - Rarely (less than a few times a month)
    - Never/don't use text messaging

A.1.2 Interview Questions

1. Do you know what social engineering is? [or phishing]
2. Have you come across SE/phishing before?
3. Where do you usually encounter social engineering attacks?
4. Where do you normally expect to encounter such an attack?
5. Who receives social engineering attacks?
6. Why might they receive them?
7. Who are these attackers?
8. What do you think attackers are trying to obtain?
9. How do these attacks trick users?
10. Why are users vulnerable to these attacks?
11. How do you think phishing attacks work (on the user's end)?
12. How do you know if you have fallen for an attack?
13. What do you usually do when you receive an attack like this?
14. How do you know if something is legit or fraudulent?
15. What do attackers do with this information?
16. On average, how often do you think you receive a phishing attack per month?
17. How many people do these attacks normally target?
18. How do these attacks spread to different individuals/devices?
19. Do you generally expect to be phished outside of [insert phishing medium previously mentioned]?
20. Have you ever seen phishing through different means? (e.g. social media, QR codes, mobile apps, SMS, website)
21. How do you think phishing attacks work for [social media/SMS/QR codes]?

A.1.3 Ethics Application: Study 1

Office of Research Ethics
CARLETON UNIVERSITY

RESEARCH INVOLVING VERY LOW RISK

This Form is for research projects meeting *all* the following criteria. If you have any doubt about whether your study may use this form, or questions about its completion, please contact the Research Ethics Office at ethics@carleton.ca or by phone to 613 520 2600 ext. 2517 (CUREB A) or ext. 4085 (CUREB B).

1. The risks to participants are very low;
2. No research procedures involve any physically invasive intervention;
3. Participants are legally capable of consenting on their own behalf, and are free from coercion or undue influence;
4. Any accidental or intentional disclosure of the participants' responses would not reasonably place participants at risk of criminal or civil liability, harmful retaliation, or be damaging to the participants' emotional or financial well-being, employability, or reputation;
5. The study does not involve recruitment by a third party aside from a paid research service such as Qualtrics or Survey Monkey; and
6. The study does not involve deception or providing incomplete information to participants.
7. This study does not primarily involve Indigenous peoples or communities.

If the project does not meet all of these conditions, then the main CUREB Protocol Form must be used.

* Please submit the Very Low Risk form as a new application in CuResearch. If this form is to replace a Release of Funds, it should be submitted as an "Event" in CuResearch under the same study file. Please see our CuResearch User Manual for directions on how to submit a new application or an event.

* Note that all of our forms are compatible with Microsoft Office. Students and staff members can download a free copy of MS Office at no charge: Students: https://carleton.ca/its/ms-offer-students/ ; Staff/Faculty: https://carleton.ca/its/all-services/computers/site-licensed-software/ms-offer-faculty/

1. Title and Date

1A Project Title

End User Mental Models of Social Engineering Attacks

1B Submission Date Date of completion of this form. Update each time the form is revised.

Click or tap to enter a date.

2. Project Team

2A Lead Researcher Last name/First name, Institutional Email, Department/Faculty and Institution if not Carleton

Kyi/Lin, Lin.Kyi@carleton.ca, School of Computer Science

- ☐ Academic or Library Staff
- ☐ Post-doctoral Fellow
- ☐ Doctoral Student
- ☒ Masters Student

If other, please describe

☐	Undergraduate Student
☐	Other

2B Academic Supervisor

☐	Same as lead researcher

Academic supervisor(s) Last name/First name, Institutional Email, Department/Faculty and Institution if not Carleton (Note, the supervisor must be copied on all correspondence with CUREB.) (Detailed instructions, Example)

Stobert/Elizabeth, Elizabeth.Stobert@carleton.ca, School of Computer Science

2C Project Team Members

☒	No other team members

List the project team members: 1) Last name/First name 2) Email address 3) Role in project 4) Department and institution (Detailed instructions, Example)

3. Project Description

3A Is this Project Funded?

☒	Yes
☐	No

If Yes, who is the award provided by, and what is the title of the award?

Carleton University Start-Up Fund (#185479)

3B Study Goal

Briefly explain the primary objective(s) of the current study.

In this study, we propose a new framework for understanding social engineering (SE) attacks. We are investigating how end users perceive SE attacks, and compare their mental models to the framework to assess gaps in user understanding. We will ask participants about their experiences with SE, and their thoughts on various aspects of SE attacks.

This current study consists of a user interview between the researcher and participants. The researcher will ask participants to fill out a demographics questionnaire, and ask questions relating to the participant's experiences and thoughts about social engineering to understand the topic.

3C **Study Rationale and Expected Benefits**

Study Rationale and Expected Benefits: Why should the study be done? What are the benefits, and to whom?

Current SE frameworks are outdated, and do not account for emerging SE attacks, such as social media phishing, QR phishing, SMS phishing, etc. Additionally, current SE research tends to focus on website and email phishing only. Therefore, there is a lack of knowledge regarding emerging SE attacks, which leaves users more vulnerable to these attacks.

The benefits are that this study will be one of the first to study and understand SE attacks beyond traditional SE attacks, and provide an updated framework for understanding SE attacks. This study will benefit security researchers, and end users by providing an update on the mental models of these less well-known attacks. Through user interviews, we will understand user mental models of the various kinds of SE attacks.

3D **Overview of Methodology and Participant Interactions**

Briefly describe the study methodology and what will be required of participants for this study

We will conduct user interviews. One researcher and one participant will meet over Zoom and discuss the participants' experiences and thoughts of social engineering. Additionally, participants will fill out a demographics questionnaire through a link sent over email.

This study will be conducted remotely, over videocall (Zoom) and screen sharing. Participants will first answer a demographics questionnaire, and then we will interview them about their knowledge of SE attacks.

We will audio record participants on Zoom, and will inform them about this. Each Zoom session will be password-protected so that no one but the researchers and participant can participate in the call.

We may ask participants if they want to do a follow-up study, which would also take place over Zoom. Participants will be asked to go through examples of SE attack prototypes and demonstrate what they would do if they were faced with these items in their own life. We will submit an additional ethics form for this study, and will get ethics approval before contacting these participants for a follow-up study.

4. Participants and Informed Consent

4A **Description of Participants**

Describe the participants and any inclusion and exclusion criteria. If using a separate sample of control participants, describe this group. (Detailed instructions, Example)

Participants will be over 18 years old, must be proficient at speaking and reading English, and must not come from a computer security

background. Participants do not need to be students, nor do they need to be in Ottawa since the study will be conducted remotely.

4B **Number of Participants (Sample size)**

How many participants will be recruited? If multiple groups of participants are involved, breakdown by participant type. Provide a justification including a statistical rationale if appropriate. (Detailed instructions, Example)

Approximately 60 participants will be recruited. We are looking to hear a wide array of perspectives from participants.

4C **Recruitment**

Describe how participants will be recruited including how contact information will be obtained. How will participants be made aware of the study, where will recruitment materials be located, and how may participants express their interest? Attach a copy of any recruitment materials including oral scripts, recruitment posters, emails and social media postings, etc.

We will post on Prolific.co (an online website for participant recruitment specifically for research studies) to recruit participants. Participants will see our posting on Prolific, and sign up for a study session (if they meet our participant criteria) using Calendly. Participants will provide their anonymous Prolific email account when signing up for a session, which only includes their Prolific ID, and not any identifiable information. We will communicate with participants using their Prolific email to send them a Zoom link and other study materials.

Please see Appendix 5 for our recruitment materials.

4D **Compensation and Remuneration**

Describe any compensation or remuneration for participants and indicate when participants will receive the compensation. What happens to compensation if a participant withdraws early?

Participants will be compensated $10 for 30 minutes of their time (paid via prolific.co).

Participants can withdraw from the study, with full compensation, up until the end of the study.

4E **Withdrawal Process**

Describe the process for a participant to withdraw their data after collection, and the time limits, if any.

Participants will have 24 hours after the interview to withdraw their consent. When participants withdraw from a study, all of their data we have collected will be deleted. This means all audio recordings and

questionnaire input will be deleted, and not used in the study analysis. Participants can anonymously contact the researcher using Prolific.co to withdraw from the study after it has completed. We will refer to participants using their Prolific ID (including the demographics and audio recordings), therefore we will be able to keep track of which materials need to be deleted based on the participant's ID.

4F Consent

Describe the process of obtaining informed consent from participants and include a copy of the consent form(s) and materials. If signed consent is not to be used, describe and justify the alternative method chosen.

Since this study will be conducted virtually, we will provide participants with an implied consent form that is incorporated in the demographics survey (see Appendix 2). At the bottom of the demographics questionnaire, it will say "By completing this survey you are agreeing to participate in this survey and be audio recorded." We are using implied consent because we want to reduce any possibility of collecting identifiable information from participants, such as signatures, which would be required in a consent form.

The researcher will be present as the participant remotely fills out the demographics form in case the participant has any questions or concerns regarding the study and consent.

4G Risks

Describe any possible physical, emotional, social, privacy, or legal risks to which participants may be exposed.

Because we are discussing mental models of SE attacks, participants may feel some shame or embarrassment if they fell for an attack, or almost fell for one. Aside from this, we do not anticipate any other forms of risk incurred by participation in the study.

4H Vulnerable Participants

Does the research project target participants from vulnerable populations (e.g., children, elderly, or prisoners), involve sensitive questions, include partial disclosure and/or mild deception, involve physical exertion, physical procedures or physical contact? If yes, justify why the project(s) still falls within the parameters of very low risk**.**

☐ Yes

☒ No

5. Data Collection, Use, and Storage

5A Collection

Describe how data will be collected and any instruments to be used. Provide a copy of any questionnaires, surveys, interview guides or other data collection materials.

Please Appendix 7 for a copy of the interview guide and demographics questionnaire to be administered to participants. We will use Qualtrics to administer the demographics questionnaire. The demographics questionnaire will include an implied consent blurb, therefore no identifiable information will be collected from consenting.

We will audio record participants during the interview, and transcribe the audio recording. Audio recordings will be stored locally on the researcher's computer. Audio recordings will be kept on a password-protected USB key, and transcriptions will be kept on a password-protected laptop. Once the audio recording is transcribed, the audio recording will be deleted.

5B Access

Aside from the PI, who will have access to research data?

Only the researchers will have access to research data.

5C Identifiability

Describe the identifiability of research data, including how codes or pseudonyms will be assigned

☐	Data will contain information that directly identifies participants.
☒	Data will contain information that may indirectly identify participants.
☐	Data will be coded with the code key stored securely and separate from identifying information.
☐	Data will be de-identified (anonymized) with any identifiers securely destroyed.

Provide any further relevant detail about the identifiability of data.

Data may contain details that could indirectly identify participants.

We will be using implied consent via the demographics form, therefore no identifiable information will be collected from the consenting process. We will use participants' Prolific IDs to link the appropriate demographics forms to the correct audio recording/transcript, therefore there is no need to keep a master list linking participant codes to names. The audio-recordings may be indirectly identifiable from the voices, but participants will be specifically instructed not to say their names on the recording. Additionally, Zoom does not record participant names or any identifying information when it audio records a meeting.

If participants have a particularly specific example of their encounter with a social engineering attack, their case may be potentially recognized by someone who is familiar with their experience. Although our study asks participants about social engineering, we will not ask participants to provide very personal accounts of their experiences with social engineering since it will not be relevant to our study. Identifying details, such as any names of organizations, will be anonymized in the transcriptions. Participants will be instructed not to state their names on the audio recording.

Audio recordings will be kept on a password-protected USB key, only accessible by the primary researcher. Audio transcriptions, and demographic questionnaire responses will be kept on a password-protected laptop only accessible by the researchers, and audio recordings will be deleted within six weeks of data collection, when audio transcripts have been created from the recordings. Zoom does not record participant names or any identifying information when it audio records a meeting.

Emails associated with sending the Zoom meeting invite and compensation will be deleted after the study session is completed and compensation has been given to the participant.

We will be using Prolific.co to recruit and communicate with participants. Prolific anonymizes participant identities, so we will not know their emails or their names, as they are referred to by their anonymous Prolific ID whenever researchers communicate with participants.

5D Data Security

Describe the physical (e.g. locked filing cabinet) and/or technical safeguards (e.g. Encryption) that will be used to securely store the collected physical and electronic data. Where will data be stored?

Data materials will be kept on a password-protected laptop, and only accessible by the researchers. Physical copies of any research materials will be kept in a locked cabinet in a locked room.

Zoom calls will be audio-recorded, and stored locally on a password-protected USB key and only the primary researcher will have access to these recordings. Once the audio recording has been transcribed, it will be deleted. Transcriptions will be kept on a password-protected laptop and will only be accessible by the researchers.

Zoom calls will be password-protected (a new Zoom room with a unique passcode will be created for each participant) so that no one but the researchers and participant can participate in the call. Data collected from Qualtrics will be stored locally on the researcher's computer. The researchers will not collect IP addresses or any kind of identifying information through Qualtrics

The demographics questionnaire will include an implied consent blurb, therefore no identifiable information will be collected from the consenting process. We will refer to participants by their anonymous Prolific ID, and their demographic questionnaire responses will be kept on a password-protected laptop to ensure only the researchers have access to these responses.

Audio recordings will be kept on a password-protected USB drive, and audio recordings will be deleted within six weeks of data collection, when audio transcripts have been created from the recordings.

Anonymized transcripts will be retained indefinitely for possible future use in research.

5E Does your research involve the use of ***personal data*** held by Carleton? If yes, you must complete a *security and confidentiality agreement* with the Carleton University Privacy Office prior to starting your research. The agreement is a contract between you and the university as to how you will manage the personal data throughout your research. If you have questions about the completion of this agreement, or best practices around privacy management for research, please contact the Carleton University Privacy Office by e-mail at *university_privacy_office@carleton.ca*.

Yes ☐ No ☒

If yes, please describe the personal information you will be collecting:

6. Attachments

6A Please indicate any attached materials.

☒	TCPS 2 Tutorial Course on Research Ethics (CORE) tutorial certificate for each team member. if requesting exemption, please justify below
☒	Sample of data collection instruments (survey questionnaires, test instruments, ect.)
☒	Supervisor approval form (if applicable)
☐	Permission letter from partner organizations (if applicable)
☐	Letters of Invitation/Information
☒	Consent forms, text or scripts
☒	Debriefing form
☒	Other, please describe in the following box

Appendix 2 - Consent form

Appendix 3 – Debriefing form

Appendix 4 – Online invitation

Appendix 5 – Supervisor approval form

Appendix 6 – Data collection instruments (demographics questionnaire, interview script, and interview questions)

Appendix 7 – TCPS2 certificates for both researchers

6B Provide a brief rationale if an attachment(s) is not available at time of submission.

7. Declarations

By submitting this form, the Lead Researcher and academic supervisor, if any, confirm that:

- The information in this Form is correct and accurately describes the research project.
- No recruitment or data collection for this protocol will start before receiving ethics clearance.
- I (we) will carry out this project in accordance with the information in this Form and the other submitted documents. No changes will be made to the research project as described in this protocol without clearance from the Research Ethics Board.
- I will promptly notify the Research Ethics Board of any ethical breaches or concerns, adverse events, unanticipated problems, protocol deviations or complaints that arise relating to this project.
- This study meets all of the conditions for eligibility listed above.

8. Comments

Do you have any comments or suggestions to improve this form?

Carleton UNIVERSITY
Canada's Capital University

Appendix 2: Informed Consent Form

Name and Contact Information of Researchers:

Lin Kyi, Carleton University, School of Computer Science
Email: Lin.Kyi@Carleton.ca

Supervisor and Contact Information:

Elizabeth Stobert, Carleton University, School of Computer Science
Email: Elizabeth.Stobert@carleton.ca

Project Title

End User Mental Models of Social Engineering Attacks

Project Sponsor and Funder (if any)

Carleton University Start-Up Fund (#185479)

Carleton University Project Clearance

Clearance #: 115226 Date of Clearance: February 11, 2021

Invitation

You are invited to take part in a research project because you are an English-speaking individual over 18 years old who does not have a computer security background. The information in this form is intended to help you understand what we are asking of you so that you can decide whether you agree to participate in this study. Your participation in this study is voluntary, and a decision not to participate will not be used against you in any way. The interview will be audio recorded during the study. As you read this form, and decide whether to participate, please ask all the questions you might have, take whatever time you need, and consult with others as you wish.

Eligibility Criteria

- *Do not come from a computer security background*
- *Are over 18 years old*
- *Are proficient at speaking and reading English*

What is the purpose of the study?

This research examines the mental models of the various types of social engineering (SE) attacks.

In this study, we propose a new framework for understanding social engineering (SE) attacks. We are investigating how end users perceive SE attacks, and analyzing their mental models to assess gaps in user understanding.

What will I be asked to do?

If you agree to take part in the study, we will ask you to:

- *Fill out a demographics questionnaire*
- *Answer some of our questions about your thoughts and experiences with social engineering attacks*

Risks and Inconveniences

We do not anticipate any risks to participating in this study.

Possible Benefits

You may not receive any direct benefit from your participation in this study. However, your participation may allow researchers to better understand how end users understand emerging social engineering attacks.

Compensation

You will be paid $10 via e-transfer (if within Canada), or given an equivalent amount through a digital Amazon gift card (if outside of Canada) for your participation in this study, which is expected to take 30 minutes.

No waiver of your rights

By signing this form, you are not waiving any rights or releasing the researchers from any liability.

Withdrawing from the study

If you withdraw your consent during the course of the study, all data collected will be deleted. This means audio recordings and questionnaire input will be deleted and not used in the study analysis. You can withdraw up to 24 hours after the interview is over. If you withdraw, you will still receive the full study compensation.

Confidentiality

We will remove all identifying information from the study data as soon as possible, which will be after this user study session.

We will treat your personal information as confidential, although absolute privacy cannot be guaranteed. No information that discloses your identity will be released or published without your specific consent. Research records may be accessed by the Carleton University Research Ethics Board in order to ensure continuing ethics compliance.

The results of this study may be published or presented at an academic conference or meeting, but the data will be presented so that it will not be possible to identify any participants.

You will be assigned a code so that your identity will not be directly associated with the data you have provided. All data, including coded information, will be kept in a password-protected file on a secure computer.

"In-session" data, such as the audio, video and chat transcript from the interview, will be stored locally on the researcher's computer. Operation data, such as meeting and performance data, will be stored and

protected by Zoom on servers located in Toronto, and Qualtrics servers located in Canada/US/EU/Australia but may be disclosed via a court order or data breach. Data collected from Zoom will be stored locally on the researcher's computer, and data collected from Qualtrics will be kept on the Qualtrics cloud. The researchers will not collect IP addresses or any kind of identifying information through Qualtrics.

We will password protect any research data that we store or transfer.

Data Retention

After the study is completed, your de-identified data will be retained for future research use.

New information during the study

In the event that any changes could affect your decision to continue participating in this study, you will be promptly informed.

Ethics review

This project was reviewed and cleared by the Carleton University Research Ethics Board B. If you have any ethical concerns with the study, please contact Carleton University Research Ethics Board (by email at ethics@carleton.ca).

Statement of consent – print and sign name

I voluntarily agree to participate in this study.	___Yes	___No
I agree to be audio recorded in this study.	___Yes	___No
I agree to be contacted for follow up research	___Yes	___No

_____ _____

Signature of participant Date

Carleton UNIVERSITY
Canada's Capital University

APPENDIX 3: DEBRIEFING

Name and Contact Information of Researchers:

Lin Kyi, Carleton University, School of Computer Science
Email: Lin.Kyi@carleton.ca

Supervisor and Contact Information:

Elizabeth Stobert, Carleton University, School of Computer Science
Email: Elizabeth.Stobert@carleton.ca

Project Title

End User Mental Models of Emerging Social Engineering Attacks

Project Sponsor and Funder (if any)

Carleton University Start-Up Fund (#185479)

Carleton University Project Clearance

Clearance #: 115226 Date of Clearance: February 11, 2021

What are we trying to learn in this research?

This research examines users' mental models of the various types of social engineering (SE) attacks. Current research tends to focus mainly on email and website phishing, but there are emerging SE attacks, such as social media phishing, SMS phishing, etc. We are trying to understand what end users understand and how they behave with these emerging SE attacks.

Why is this important to scientists or the general public?

The research on social engineering is outdated and has not yet looked at mental models of newer SE attacks. Therefore, many end users are left vulnerable to these newer SE attacks because there is less user awareness and research on them. We hope to bridge the gap between current research and what is happening in the real world by providing one of the first studies understanding these emerging attacks.

What are our hypotheses and predictions?

We predict that end users will be less aware about emerging SE attacks, and will not be as knowledgeable about what to do when they encounter these attacks. We expect that there will be a large gap in understanding of SE attacks between end users and our framework of social engineering.

Where can I learn more?

There are several online resources available where you can learn more about social engineering. The following resources can help you identify social engineering attacks, and learn how to can avoid falling victim to these attacks:

https://www.social-engineer.org/about/

https://digitalguardian.com/blog/dont-get-hooked-how-recognize-and-avoid-phishing-attacks-infographic

https://www.social-engineer.org/social-engineering/social-engineering-infographic/

https://us.norton.com/internetsecurity-emerging-threats-what-is-social-engineering.html#:~:text=Social%20engineering%20is%20the%20act,taking%20action%2C%20usually%20through%20technology.&text=A%20social%20engineer%2C%20though%2C%20could,into%20divulging%20their%20login%20credentials.

What if I have questions later?

If you have any remaining concerns, questions, or comments about the experiment, please feel free to contact Lin Kyi (Principal Investigator), at Lin.Kyi@carleton.ca, or Dr. Elizabeth Stobert (Faculty Sponsor), at Elizabeth.Stobert@carleton.ca.

If you have any ethical concerns with the study, please contact the Carleton University Research Ethics Board-B or via email at ethics@carleton.ca).

Thank you for participating in this research!

Appendix 4: Online Invitation

To be posted on Prolific.co:

Volunteers needed for a study about social engineering

We are looking for volunteers for a social engineering study. The project aims to understand how everyday users understand and interact with various kinds of social engineering attacks. The study takes place online through a Zoom video call.

You will be asked to fill out a demographics questionnaire and participate in a 30-minute interview. The interview will be audio recorded.

To be eligible, you must be over 18 years old, speak English, and not come from a computer security background.

The study will take place remotely, over a Zoom video call and should take approximately 30 minutes to complete. Participants will receive CAD$10 (paid through Prolific).

If you are interested, please follow the Calendly link to book your interview session.

This research has been cleared by Carleton University Research Ethics Board B Clearance #115226

Should you have any ethical concerns with the study, please contact the REB Chair, Carleton University Research Ethics Board-B (by email: ethics@carleton.ca). For all other questions about the study, please contact the researcher.

Carleton UNIVERSITY
Canada's Capital University

Appendix 6: Supervisor signature

SUPERVISOR/SPONSOR SIGNATURE FORM

For consideration of submitted ethics protocols, the Carleton University Research Ethics Boards require evidence that all student protocol documents (i.e., undergraduate, graduate and post-doctoral fellows) have been reviewed and approved by a faculty supervisor or sponsor.

Instructions:

After filling out the details in the text below, faculty sponsor/supervisors should either

1) Print and scan this document, or
2) Email the text of the document (below; with signature optional) to the lead researcher

The form or email can then be uploaded (in PDF format) with the protocol, to the CUResearch.

As the faculty supervisor or sponsor, I confirm that I have reviewed and that I approve for submission for ethics review, the protocol entitled "End User Mental Models of Social Engineering Attacks" from Lin Kyi on January 19, 2021.

Signature of Faculty Supervisor/Sponsor: Elizabeth Stobert

Name of Faculty Supervisor/Sponsor: Elizabeth Stobert

Date: January 19, 2021

Appendix 7: Data Collection Instruments

Hi, my name is Lin and I'm currently a Master's student at Carleton. In this study, I am trying to understand how everyday users understand and interact with social engineering attacks.

I've sent a link in the Zoom chat. Could you please open it, and fill it out. There is the demographics form. By completing this survey, you are agreeing to participate in the study and be audio recorded. Let me know if you have any questions.

This study should take around 30 minutes. This study will be in the form of an interview – I will ask some questions about the topic to see what general computer users know. We aren't testing your knowledge, but rather to see what users know so we can come up with better solutions to help users.

During the study, we will ask that you not provide your name or any identifiable information so that we can maintain your privacy.

DEMOGRAPHICS SURVEY:

1. What is your gender identity?
    a. Male
    b. Female
    c. Non-binary / third gender
    d. Prefer not to say
2. How old are you?
3. How knowledgeable are you about computer security?
    a. Not knowledgeable at all
    b. Slightly knowledgeable
    c. Moderately knowledgeable
    d. Very knowledgeable
    e. Extremely knowledgeable
4. What is your highest level of education?
    a. High school
    b. Community college
    c. Trades
    d. Undergraduate degree
    e. Graduate or professional degree
5. What is/was your area of study?
6. What is your occupation?
7. How susceptible do you think you are to social engineering/phishing?
    a. Extremely susceptible
    b. Very susceptible
    c. Moderately susceptible
    d. Slightly susceptible
    e. Not susceptible at all

8. What kind of smartphone do you own?
    a. iPhone (iOS)
    b. Android
    c. Both
    d. Other: please specify
    e. Don't own a smartphone
9. What type of computer operating system are you familiar with? [select all that apply]
    a. Macbook/Apple computer
    b. Microsoft operating system
    c. Linux operating system
    d. Other: please specify:
    e. Don't own a computer
10. How often do you use the following technologies:
    a. Email
        i. Daily
        ii. A few times a week
        iii. Rarely (less than a few times a month)
        iv. Never/Don't have
    b. Social media
        i. Daily
        ii. A few times a week
        iii. Rarely (less than a few times a month)
        iv. Never/Don't have
    c. Downloading and using mobile apps
        i. Daily
        ii. A few times a week
        iii. Rarely (less than a few times a month)
        iv. Never/Don't have
    d. Text messaging/SMS
        i. Daily
        ii. A few times a week
        iii. Rarely (less than a few times a month)
        iv. Never/Don't have
    e. QR codes
        i. Daily
        ii. A few times a week
        iii. Rarely (less than a few times a month)
        iv. Never/Don't have
11. By completing this questionnaire, you are agreeing to participate in this study and be audio recorded.

USER INTERVIEWS:

1. Attack Formulation Stage

    a. Do you know what social engineering is? [or phishing]
    b. Have you come across SE/phishing before?
    c. Medium:
        i. Where do people usually encounter SE attacks?
        ii. Where do you normally expect to encounter such an attack?
    d. Target:
        i. Who receives SE attacks?
        ii. Why might they receive them?
    e. Goal:
        i. Who are these attackers?
        ii. What do you think attackers are trying to obtain?
2. Persuasive Techniques
    a. How do these attacks trick users?
    b. Why are users vulnerable to these attacks?
    c. What emotions or feelings do these attacks elicit?
        i. Does it vary by the type of attack?
3. Conduct the Attack
    a. How do you think phishing attacks work?
    b. How do you know if you have fallen for an attack?
    c. What do you usually do when you receive an attack like this?
    d. How do you know if something is legit or fraudulent?
4. Utilize the Information
    a. What do attackers do with this information?
    b. On average, how often do you think you receive a phishing attack per month? [phishing, social media, phone, etc.]
5. Repeat
    a. How many people do these attacks normally target?
    b. How do these attacks spread to different individuals/devices?
6. Other attacks
    a. Have you ever seen phishing through different means?
        i. E.g. social media, QR codes, mobile apps, SMS, website
    b. How do you think phishing attacks work for [social media/SMS/QR codes]?
    c. Do you generally expect to be phished through these means? [or safe?]

This is the end of our user study. Thank you so much for your participation!

Do you have any questions for me?

I have sent a debriefing form to you over email so you can take a look at more specific details regarding this study, and it includes my contact information in case you do have any additional questions or concerns.

A.2 Study 2

A.2.1 Survey

Demographics

1. What is your gender identity?
    - Male
    - Female
    - Non-binary
    - Prefer not to say
2. How old are you?
    - 18–25
    - 26–35
    - 36–45
    - 46–55
    - 56–65
    - 65+
    - Prefer not to say
3. How knowledgeable are you about computer security?
    - Not knowledgeable at all
    - Slightly knowledgeable
    - Moderately knowledgeable
    - Very knowledgeable
    - Extremely knowledgeable
4. What is your current level of education? If completed, please select your highest level of education.
    - High school
    - Community college
    - Trades or vocational training
    - Undergraduate degree
    - Graduate or professional degree
    - No school completed
5. How susceptible do you think you are to phishing (social engineering)?
    - Extremely susceptible

- Very susceptible
- Moderately susceptible
- Slightly susceptible
- Not susceptible at all

6. How often do you use email?
   - Daily
   - Sometimes (a few times a week)
   - Rarely (less than a few times a month)
   - Never/don't have email
7. How often do you use social media?
   - Daily
   - Sometimes (a few times a week)
   - Rarely (less than a few times a month)
   - Never/don't have social media
8. How often do you use text messaging (or similar) on your phone?
   - Daily
   - Sometimes (a few times a week)
   - Rarely (less than a few times a month)
   - Never/don't use SMS
9. Have you ever received any training, whether formally or on your own time, to identify phishing attacks?
   - Yes
   - No
10. (If participant selected "Yes" to Q.9) Where did you receive this training?
    - At the workplace
    - Through a formal course
    - Looked it up in my free time or out of my own interest
11. (If participant selected "Yes" to Q.9) When did you most recently receive this training?
    - Within the past 6 months
    - Last 5 years
    - 5+ years ago

12. Have you ever been the victim of a phishing attack?
    - Yes
    - No
    - Unsure

Participant Evaluations of Social Engineering Attacks

All participants were shown a screenshot of an social engineering attack example, and asked the following questions for each example:

1. Have you ever seen a phishing attack like this before?
    - Yes
    - No
2. What would you do if you encountered this item?
    - Click the link and/or call the number to investigate
    - Delete or ignore it
    - Verify or Google to see if it is legitimate
    - Block or report the sender
    - Change the password to your account
3. How confident are you that your actions will prevent you from potential consequences?
    - 1: Not confident
    - 2: Slightly confident
    - 3: Moderately confident
    - 4: Confident
    - 5: Very confident
4. Who is this item most likely being sent to?
    - Anyone
    - Specific group based on shared interests, demographics, etc.
    - Only the one recipient
5. How confident do you feel about your response to the question above?
    - 1: Not confident
    - 2: Slightly confident

- 3: Moderately confident
- 4: Confident
- 5: Very confident

6. Which psychological technique is primarily being used to persuade the reader that this item is legitimate?

   - Fear
   - Greed
   - Giving the reader a fake story
   - Curiosity

7. How confident do you feel about your response to the question above?

   - 1: Not confident
   - 2: Slightly confident
   - 3: Moderately confident
   - 4: Confident
   - 5: Very confident

8. What what point would you be in most danger if you interacted with this item?

   - When you receive the item
   - When you click any links or call the phone number
   - When you enter personal information
   - When you purchase the item or pay a fee
   - When you download something from the attack item (*e.g.,* download an app, file)

9. How confident do you feel about your response to the question above?

   - 1: Not confident
   - 2: Slightly confident
   - 3: Moderately confident
   - 4: Confident
   - 5: Very confident

10. What is the attack trying to gain from victims?

    - Money
    - Personal information
    - Hack into an account or device

11. How confident do you feel about your response to the question above?
    - 1: Not confident
    - 2: Slightly confident
    - 3: Moderately confident
    - 4: Confident
    - 5: Very confident
12. Do you think this attack is generalized, or targeted?
    - Generalized (sent to many individuals, regardless of interests/demographics)
    - Targeted (sent to Only the recipient, or a smaller group based on interests/demograohics)
13. How confident do you feel about your response to the question above?
    - 1: Not confident
    - 2: Slightly confident
    - 3: Moderately confident
    - 4: Confident
    - 5: Very confident
14. If you saw this in real life, how likely would you be to believe it was legitimate?
    - 1: Not likely
    - 2: Slightly likely
    - 3: Moderately likely
    - 4: Likely
    - 5: Very likely
15. Which factors did you consider for your decision? (Select all that aply)
    - Spelling and grammar
    - Design of the item (graphics, logos, fonts, etc.)
    - Offer being presented
    - Sender or organization being impersonated
    - Presence/absence of a link or call to action
    - Other:
16. How vulnerable do you think others are to falling for this kind of attack?
    - 1: Not likely
    - 2: Slightly likely

- 3: Moderately likely
- 4: Likely
- 5: Very likely

A.2.2 Table of Attack Examples

Table A.1: The 21 social engineering attack examples shown to participants.

Attack Description	**Attack Vector**	**Target Audience**	**Persuasive Technique(s)**	**Call to Action**	**Attack Goal**	**Attack Scalability**
Impersonating Koodo Mobile, offering to help if users call the specified number	SMS	Anyone	Greed, curiosity	Giving information	Personal information	Generalized
Recipient has an unpaid postage cost for an attempted delivery	SMS	Anyone	Fear	Paying a fee	Financial gain	Generalized
Selling various kinds of educational software for discounted prices	SMS	Targeted	Greed, curiosity	Paying a fee	Financial gain	Generalized
Impersonating HSBC, claiming the user has a new payee request to Peter Black. If the user does not know Peter Black, they must click the link presented	SMS	Anyone	Fear, curiosity	Giving banking information	Personal information	Generalized

Continued on next page

Table A.1 – *Continued from previous page*

Attack Description	**Attack Vector**	**Target Audience**	**Persuasive Technique(s)**	**Call to Action**	**Attack Goal**	**Attack Scalability**
Impersonating a user's daughter and asking the mother for emergency money	SMS	Only the recipient	Pretexting	Paying a fee	Financial gain	Targeted
Impersonating a tax collection agency, offering help and claiming the user may be eligible for a large tax refund	SMS	Anyone	Greed, curiosity	Giving personal information	Personal information	Generalized
Recipient must reschedule a new delivery and pay a delivery fee	Email	Anyone	Fear	Paying a fee	Financial gain	Generalized
Impersonating a rich, dying man who will give the recipient money in exchange for a small fee	Email	Anyone	Greed, pretexting	Paying a fee	Financial gain	Generalized
Impersonating Canada Post asking users to fill out a survey in exchange for a cash prize	Email	Anyone	Greed	Giving personal information	Personal information	Generalized
Impersonating Instagram telling the user to reset their password	Email	Anyone	Fear	Give login credentials	Hack into an account	Generalized

Continued on next page

Table A.1 – *Continued from previous page*

Attack Description	**Attack Vector**	**Target Audience**	**Persuasive Technique(s)**	**Call to Action**	**Attack Goal**	**Attack Scalability**
Landlord of non-existent, affordable apartment asking potential renters to sign a form and send a deposit to secure the apartment	Email	Targeted	Greed	Paying the deposit	Financial gain	Generalized
YouTube ad claiming users can get a free $100 gift card if they click a link to watch a video	Ad	Targeted	Greed, curiosity	Clicking the link	Install malware	Targeted
Facebook post advertising discounted pet beds	Ad	Anyone	Greed	Paying for the item	Financial Gain	Generalized
Kijiji post for skilled labourers who can gain access to job opportunities by downloading an app	Ad	Targeted	Greed	Download an app	Install malware	Generalized
Facebook ad impersonating Walmart advertising iPads for sale for $1.45	Ad	Targeted	Greed	Buying the item	Financial gain	Generalized

Continued on next page

Table A.1 – *Continued from previous page*

Attack Description	**Attack Vector**	**Target Audience**	**Persuasive Technique(s)**	**Call to Action**	**Attack Goal**	**Attack Scalability**
Facebook ad selling unclaimed Amazon items	Ad	Targeted	Curiosity, greed	Buying the item	Financial gain	Generalized
Tweet impersonating Norton Security selling an anti-virus software	Social media	Anyone	Fear	Buying the item	Financial gain	Generalized
Discord direct message trying to steal tokens from users	Social media	Targeted	Curiosity, greed	Giving information	Financial gain	Generalized
Facebook post impersonating Unicef advertising student internships	Social media	Targeted	Greed	Giving personal information	Personal information	Generalized
Facebook direct message from a friend whose account has been hacked	Social media	Only the recipient	Curiosity, greed	Giving personal information	Personal information	Targeted
Instagram direct message from an account claiming they can make your Instagram account verified	Social media	Targeted	Greed	Paying a fee	Financial information	Targeted

A.2.3 Ethics Application: Study 2

Office of Research Ethics
CARLETON UNIVERSITY

RESEARCH INVOLVING VERY LOW RISK

This Form is for research projects meeting *all* the following criteria. If you have any doubt about whether your study may use this form, or questions about its completion, please contact the Research Ethics Office at ethics@carleton.ca or by phone to 613 520 2600 ext. 2517 (CUREB A) or ext. 4085 (CUREB B).

1. The risks to participants are very low;
2. No research procedures involve any physically invasive intervention;
3. Participants are legally capable of consenting on their own behalf, and are free from coercion or undue influence;
4. Any accidental or intentional disclosure of the participants' responses would not reasonably place participants at risk of criminal or civil liability, harmful retaliation, or be damaging to the participants' emotional or financial well-being, employability, or reputation;
5. The study does not involve recruitment by a third party aside from a paid research service such as Qualtrics or Survey Monkey; and
6. The study does not involve deception or providing incomplete information to participants.
7. This study does not primarily involve Indigenous peoples or communities.

If the project does not meet all of these conditions, then the main CUREB Protocol Form must be used.

* Please submit the Very Low Risk form as a new application in CuResearch. If this form is to replace a Release of Funds, it should be submitted as an "Event" in CuResearch under the same study file. Please see our CuResearch User Manual for directions on how to submit a new application or an event.

* Note that all of our forms are compatible with Microsoft Office. Students and staff members can download a free copy of MS Office at no charge: Students: https://carleton.ca/its/ms-offer-students/ ; Staff/Faculty: https://carleton.ca/its/all-services/computers/site-licensed-software/ms-offer-faculty/

1. Title and Date

1A Project Title

Evaluating User Confidence and Accuracy for Social Engineering (SE) Attacks

1B Submission Date Date of completion of this form. Update each time the form is revised.

5/12/2021

2. Project Team

2A Lead Researcher Last name/First name, Institutional Email, Department/Faculty and Institution if not Carleton

Kyi/Lin, Lin.Kyi@carleton.ca, School of Computer Science

- ☐ Academic or Library Staff
- ☐ Post-doctoral Fellow
- ☐ Doctoral Student
- ☒ Masters Student

If other, please describe

☐	Undergraduate Student
☐	Other

2B Academic Supervisor

Academic supervisor(s) Last name/First name, Institutional Email, Department/Faculty and Institution if not Carleton (Note, the supervisor must be copied on all correspondence with CUREB.) (Detailed instructions, Example)

☐	Same as lead researcher

Stobert/Elizabeth, Elizabeth.Stobert@carleton.ca, School of Computer Science

2C Project Team Members

List the project team members: 1) Last name/First name 2) Email address 3) Role in project 4) Department and institution (Detailed instructions, Example)

☒	No other team members

3. Project Description

3A Is this Project Funded?

If Yes, who is the award provided by, and what is the title of the award?

☒	Yes
☐	No

Carleton University Start-Up Fund (#185479)

3B Study Goal

Briefly explain the primary objective(s) of the current study.

We will use a survey to understand how end users evaluate various kinds of SE attack examples. We are investigating how users evaluate SE attacks, and plan to compare it to how accurate they are in their evaluations to assess gaps in user understanding.

3C Study Rationale and Expected Benefits

Study Rationale and Expected Benefits: Why should the study be done? What are the benefits, and to whom?

There is a lack of knowledge regarding mental models of SE attacks, which leaves users more vulnerable because current solutions are not adequate. We are expanding the literature on user mental models of SE attacks by looking at how users perceive and evaluate SE attacks of

the entire attack cycle, such as how they think they are targeted for these attacks, how they are spread, and when they are activated.

This study will benefit security researchers and end users by providing an update on the mental models of the SE attack cycle. Additionally, this study will help us come up with more effective solutions for end users to prevent them from falling for SE attacks and detecting them with better accuracy.

3D **Overview of Methodology and Participant Interactions**

Briefly describe the study methodology and what will be required of participants for this study

We will be using a survey for this study. First, participants will fill out a demographics questionnaire, and then will evaluate various screenshots of SE attack examples. We will be recruiting participants through Prolific.co, an online crowdsourcing platform specifically for research studies.

Participants will be asked to view screenshots of example SE attacks and respond to a series of survey questions asking about how they understand various parts of the sample attacks. Sample attacks are in Appendix 5 (Data collection materials). We will be providing multiple-choice responses to reduce the amount of personal information participants may give us. The SE attack examples are screenshots of real or fabricated SE attacks, so participants are not directly interacting with these attacks. We will inform participants to not type in any of the links we present, inform them these are SE attacks to keep them protected, and will not ask participants to provide any personal information in our study. We will remind participants that we are not testing their knowledge, but rather want to see how they evaluate these various attack examples.

We will not ask participants for a follow-up study. This study should take around 20 minutes, and will be done completely online.

To reduce participant fatigue, we will divide up participants so each participant only sees 10 examples out of the possible 30 examples. We plan to have every participant see a different group of 10 examples.

4. Participants and Informed Consent

4A **Description of Participants**

Describe the participants and any inclusion and exclusion criteria. If using a separate sample of control participants, describe this group. (Detailed instructions, Example)

Participants will be over 18 years old, must be proficient at speaking and reading English, and must not come from a computer security background. Participants do not need to be students, nor do they need to be in Ottawa since the study will be conducted remotely.

4B **Number of Participants (Sample size)**

How many participants will be recruited? If multiple groups of participants are involved, breakdown by participant type. Provide a justification including a statistical rationale if appropriate. (Detailed instructions, Example)

200

4C **Recruitment**

Describe how participants will be recruited including how contact information will be obtained. How will participants be made aware of the study, where will recruitment materials be located, and how may participants express their interest? Attach a copy of any recruitment materials including oral scripts, recruitment posters, emails and social media postings, etc.

We will share a post on Prolific.co (an online website for participant recruitment specifically for research studies) to recruit participants. Participants will see our posting on Prolific, and can participate in our study. Our Prolific post will link to a Qualtrics survey; Prolific is only used to recruit and pay participants.

Prolific ensures privacy and confidentiality for participants, therefore we will not know the identities of our participants.

Please see Appendix 5 for our recruitment materials.

4D **Compensation and Remuneration**

Describe any compensation or remuneration for participants and indicate when participants will receive the compensation. What happens to compensation if a participant withdraws early?

Participants will be compensated CAD $3 for 20 minutes of their time (paid via prolific.co).

Participants will have 24 hours after the survey to withdraw their consent, without compensation. When participants withdraw from a study, all of their data we have collected will be deleted.

To withdraw during the study, participants only need to exit the survey; we will take their incomplete survey as a sign they withdrew. To withdraw from the study after completion, participants will need to message the researcher through the Prolific platform, or by contacting Lin at Lin.Kyi@carleton.ca using their Prolific email address. This ensures that participants' identities are kept anonymous. The researcher's contact information will be available in the consent form.

4E **Withdrawal Process**

Describe the process for a participant to withdraw their data after collection, and the time limits, if any.

Participants will have up to 24 hours after the interview to withdraw their consent. Participants will not be compensated if they withdraw from the study, therefore we will pay participants 24 hours after they participate in the study. To withdraw during the study, participants only need to exit the survey; we will take their incomplete survey as a sign they withdrew. To withdraw from the study after completion, participants will need to message the research through Prolific's messaging platform, or email the lead researcher at Lin.Kyi@carleton.ca. This ensures that participants' identities are kept anonymous.

When participants withdraw from a study, all of their data we have collected will be deleted. Participants can anonymously contact the researcher using Prolific.co to withdraw from the study after it has completed. We will refer to participants using their Prolific ID (including the demographics and audio recordings), therefore we will be able to keep track of which materials need to be deleted based on the participant's ID.

4F Consent

Describe the process of obtaining informed consent from participants and include a copy of the consent form(s) and materials. If signed consent is not to be used, describe and justify the alternative method chosen.

Since this study will be conducted virtually, we will provide participants with an implied consent form that is incorporated in the survey (see Appendix 2). At the bottom of the survey, it will say "By completing this survey you are agreeing to participate in this survey." We are using implied consent because we want to reduce any possibility of collecting identifiable information from participants, such as signatures, which would be required in a consent form.

Participants can anonymously contact the researcher through the Prolific platform if they have any questions about the consent process.

4G Risks

Describe any possible physical, emotional, social, privacy, or legal risks to which participants may be exposed.

Because we are studying mental models of SE attacks, participants may feel some shame or embarrassment if they fell for an attack, or almost fell for one.

All SE attack examples in this study are screenshots of real or fabricated attacks; participants will be informed these are screenshots, and will be asked to not type in any of the links they see or disclose any personal information in the study. Participants will not be directly interacting with these attacks, only viewing the screenshots and answering our multiple choice questions to evaluate these examples.

4H **Vulnerable Participants**

Does the research project target participants from vulnerable populations (e.g., children, elderly, or prisoners), involve sensitive questions, include partial disclosure and/or mild deception, involve physical exertion, physical procedures or physical contact? If yes, justify why the project(s) still falls within the parameters of very low risk**.**

☐ Yes

☒ No

5. Data Collection, Use, and Storage

5A **Collection**

Describe how data will be collected and any instruments to be used. Provide a copy of any questionnaires, surveys, interview guides or other data collection materials.

Please see Appendix 7 for a copy of the survey which will be administered to participants. We will use Qualtrics to administer the survey.

5B **Access**

Aside from the PI, who will have access to research data?

Only the researchers will have access to research data.

5C **Identifiability**

Describe the identifiability of research data, including how codes or pseudonyms will be assigned

☐	Data will contain information that directly identifies participants.
☐	Data will contain information that may indirectly identify participants.
☐	Data will be coded with the code key stored securely and separate from identifying information.
☒	Data will be de-identified (anonymized) with any identifiers securely destroyed.

Provide any further relevant detail about the identifiability of data.

5D **Data Security** Describe the physical (e.g. locked filing cabinet) and/or technical safeguards (e.g. Encryption) that will be used to securely store the collected physical and electronic data. Where will data be stored?

Data materials will be kept on a password-protected laptop, and only accessible by the researchers. Physical copies of any research materials will be kept in a locked cabinet in a locked room.

Data collected from Qualtrics will be stored locally on the researcher's computer. The researchers will not collect IP addresses or any kind of identifying information through Qualtrics

The survey will include an implied consent blurb, therefore no identifiable information will be collected from the consenting process. We will refer to participants by their anonymous Prolific ID, and their demographic questionnaire responses will be kept on a password-protected laptop to ensure only the researchers have access to these responses.

Data from this study will be kept indefinitely for use in future research projects.

5E Does your research involve the use of ***personal data*** held by Carleton? If yes, you must complete a *security and confidentiality agreement* with the Carleton University Privacy Office prior to starting your research. The agreement is a contract between you and the university as to how you will manage the personal data throughout your research. If you have questions about the completion of this agreement, or best practices around privacy management for research, please contact the Carleton University Privacy Office by e-mail at *university_privacy_office@carleton.ca*.

Yes ☐ No ☒

If yes, please describe the personal information you will be collecting:

6. Attachments

6A Please indicate any attached materials.

☒	TCPS 2 Tutorial Course on Research Ethics (CORE) tutorial certificate for each team member. if requesting exemption, please justify below
☒	Sample of data collection instruments (survey questionnaires, test instruments, ect.)
☐	Supervisor approval form (if applicable)
☐	Permission letter from partner organizations (if applicable)
☒	Letters of Invitation/Information

☒	Consent forms, text or scripts
☒	Debriefing form
☐	Other, please describe in the following box

Appendix 2 - Consent form
Appendix 3 – Debriefing form
Appendix 4 – Online invitation
Appendix 5 – Data collection instruments (demographics questionnaire, survey)
Appendix 6 – TCPS2 certificates for both researchers

6B Provide a brief rationale if an attachment(s) is not available at time of submission.

7. Declarations

By submitting this form, the Lead Researcher and academic supervisor, if any, confirm that:

- The information in this Form is correct and accurately describes the research project.
- No recruitment or data collection for this protocol will start before receiving ethics clearance.
- I (we) will carry out this project in accordance with the information in this Form and the other submitted documents. No changes will be made to the research project as described in this protocol without clearance from the Research Ethics Board.
- I will promptly notify the Research Ethics Board of any ethical breaches or concerns, adverse events, unanticipated problems, protocol deviations or complaints that arise relating to this project.
- This study meets all of the conditions for eligibility listed above.

8. Comments

Do you have any comments or suggestions to improve this form?

Carleton UNIVERSITY
Canada's Capital University

Appendix 2: Informed Consent Form

Name and Contact Information of Researchers:

Lin Kyi, Carleton University, School of Computer Science
Email: Lin.Kyi@carleton.ca

Supervisor and Contact Information:

Elizabeth Stobert, Carleton University, School of Computer Science
Email: Elizabeth.Stobert@carleton.ca

Project Title

Evaluating User Confidence and Accuracy for Social Engineering (SE) Attacks

Project Sponsor and Funder (if any)

Carleton University Start-Up Fund (#185479)

Carleton University Project Clearance

Clearance #: 115762 Date of Clearance: May 28, 2021

Invitation

You are invited to take part in a research project because you are an English-speaking adult with no computer security background. The information in this form is intended to help you understand what we are asking of you so that you can decide whether you agree to participate in this study. Your participation in this study is voluntary, and a decision not to participate will not be used against you in any way. As you read this form, and decide whether to participate, please ask all the questions you might have, take whatever time you need, and consult with others as you wish.

What is the purpose of the study?

This study aims to understand how end users evaluate various kinds of social engineering (SE) attacks. Social engineering, also commonly known as phishing, is defined as the use of psychological manipulation as a method of compromising computer security systems. We are expanding the literature on user mental models of SE attacks by looking at how users perceive and evaluate SE attacks of the entire attack cycle, such as how they think they are targeted for these attacks, how they are spread, and when they are activated.

What will I be asked to do?

If you agree to take part in the study, we will ask you to:

- Fill out an online demographics survey and evaluate various social engineering examples
    - These attack examples are screenshots of attacks, and you will not be asked to directly interact with them
- This is expected to take under 20 minutes, and is done completely online

Risks and Inconveniences

We do not anticipate any risks to participating in this study.

Possible Benefits

You may be more aware of the various kinds of emerging social engineering attacks that exist, and understand how to identify them through this research. Additionally, your participation may allow researchers to better understand how end users understand social engineering attacks.

Compensation/Incentives

You will be paid CAD $3 in exchange for approximately 20 minutes of your time.

No waiver of your rights

By signing this form, you are not waiving any rights or releasing the researchers from any liability.

Withdrawing from the study

If you withdraw your consent during the course of the study, all data collected will be deleted. This means all questionnaire input will be deleted and not used in the study analysis. To withdraw during the study, you only need to exit the survey. You can withdraw up to 24 hours after the interview is over by contacting the researcher at Lin.Kyi@carleton.ca; please email using your Prolific email so we can keep track of which response to delete. Alternatively, you may message the researcher through the Prolific platform. If you withdraw from the survey, you will not receive compensation.

Confidentiality

We will remove all identifying information from the study data as soon as possible, which will be after this user study session.

We will treat your personal information as confidential, although absolute privacy cannot be guaranteed. No information that discloses your identity will be released or published without your specific consent. Research records may be accessed by the Carleton University Research Ethics Board in order to ensure continuing ethics compliance.

The results of this study may be published or presented at an academic conference or meeting, but the data will be presented so that it will not be possible to identify any participants.

You will be assigned a code so that your identity will not be directly associated with the data you have provided. All data, including coded information, will be kept in a password-protected file on a secure computer.

Operation data, such as meeting and performance data, will be stored and protected by Qualtrics servers located in Canada/US/EU/Australia but may be disclosed via a court order or data breach. Data collected from Qualtrics will be kept on the Qualtrics cloud. The researchers will not collect IP addresses or any kind of identifying information through Qualtrics.

We will password protect any research data that we store or transfer.

Data Retention

After the study is completed, your de-identified data will be retained for future research use.

New information during the study

In the event that any changes could affect your decision to continue participating in this study, you will be promptly informed.

Ethics review

This project was reviewed and cleared by the Carleton University Research Ethics Board B. If you have any ethical concerns with the study, please contact Carleton University Research Ethics Board (by email at ethics@carleton.ca).

Statement of consent – print and sign name

I voluntarily agree to participate in this study. ___Yes ___No

_____ _____

Signature of participant Date

Carleton UNIVERSITY
Canada's Capital University

APPENDIX 3: DEBRIEFING TEMPLATE

Name and Contact Information of Researchers:

Lin Kyi, Carleton University, School of Computer Science
Email: Lin.Kyi@carleton.ca

Supervisor and Contact Information:

Elizabeth Stobert, Carleton University, School of Computer Science
Email: Elizabeth.Stobert@carleton.ca

Project Title

Evaluating User Confidence and Accuracy for Social Engineering (SE) Attacks

Project Sponsor and Funder (if any)

Carleton University Start-Up Fund (#185479)

Carleton University Project Clearance

Clearance #: 115762 Date of Clearance: May 28, 2021

What are we trying to learn in this research?

This study aims to understand how end users evaluate various kinds of SE attacks. We are investigating how users evaluate SE attacks, and compare it to how accurate they are in their evaluations to assess gaps in user understanding.

Why is this important to scientists or the general public?

The research on social engineering is outdated and has not yet looked at mental models of the entire SE attack cycle. Therefore, many end users are still left vulnerable to SE attacks because there is less user awareness and research. We hope to bridge the gap between current research and what is happening in the real world by providing one of the first studies understanding mental models of the SE attack cycle.

What are our hypotheses and predictions?

We believe there are gaps in user confidence and actual accuracy in evaluating various SE attacks. Specific parts of the SE framework, such as attack scalability, and the call to action may produce wider gaps in user confidence and accuracy compared to other aspects of the framework such as attack medium and audience.

Where can I learn more?

There are several online resources available where you can learn more about social engineering. The following resources can help you identify social engineering attacks, and learn how to can avoid falling victim to these attacks:

https://www.social-engineer.org/about/

https://digitalguardian.com/blog/dont-get-hooked-how-recognize-and-avoid-phishing-attacks-infographic

https://www.social-engineer.org/social-engineering/social-engineering-infographic/

https://us.norton.com/internetsecurity-emerging-threats-what-is-social-engineering.html#:~:text=Social%20engineering%20is%20the%20act,taking%20action%2C%20usually%20through%20technology.&text=A%20social%20engineer%2C%20though%2C%20could,into%20divulging%20their%20login%20credentials.

What if I have questions later?

If you have any remaining concerns, questions, or comments about the experiment, please feel free to contact Lin Kyi (Principal Investigator), at Lin.Kyi@carleton.ca, or Dr. Elizabeth Stobert (Faculty Sponsor), at Elizabeth.Stobert@carleton.ca.

If you have any ethical concerns with the study, please contact the Carleton University Research Ethics Board-B or via email at ethics@carleton.ca).

Thank you for participating in this research!

Carleton UNIVERSITY
Canada's Capital University

Appendix 4 - **Online Invitation**

To be posted on Prolific.co.

Volunteers needed for a study about social engineering

We are looking for volunteers for a social engineering study. This project aims to have participants evaluate several social engineering examples. Participants will receive CAD $3 in exchange for 20 minutes. The study takes place at Qualtrics.com, and is a survey.

The study aims to better understand end user evaluations of social engineering attack examples. Social engineering, also commonly known as phishing, is defined as the use of psychological manipulation as a method of compromising computer security systems. You will be asked to fill out a demographics questionnaire and fill out our survey about social engineering attacks.

To be eligible, you must be over 18 years old, speak English, and not come from a computer security background.

The study will take place on Qualtrics.com and should take approximately 20 minutes to complete. You will be compensated $3 for your participation.

This research has been cleared by Carleton University Research Ethics Board B Clearance #115762.

Should you have any ethical concerns with the study, please contact the REB Chair, Carleton University Research Ethics Board-B (by email: ethics@carleton.ca). For all other questions about the study, please contact the researcher.

Appendix 5: Data collection instruments

Consent:

- By checking this box, I am consenting to participate in this study

Demographics:

- Please enter your Prolific ID

- What is your gender identity?
  - Man
  - Woman
  - Non-binary/third gender
  - Other gender identity:
  - Prefer not to say

- How old are you?
  - 18 – 25
  - 26 – 35
  - 36 – 45
  - 46 – 55
  - 56 – 65
  - 65+

- How knowledgeable are you about computer security?
  - Not knowledgeable at all
  - Slightly knowledgeable
  - Moderately knowledgeable
  - Knowledgeable
  - Very knowledgeable

- What is your highest level of education?
  - High school
  - Community college
  - University (Undergraduate degree)
  - Graduate or Professional school
  - Trades

- How likely do you think it is that you will be the victim of a social engineering (phishing) attack?
  - Extremely likely
  - Very likely
  - Moderately likely

    - Slightly likely
    - Not likely at all

- How often do you use email?
    - Everyday
    - Sometimes (a few times a week)
    - Rarely (less than a few times a month)
    - Never/don't have

- How often do you use social media?
    - Everyday
    - Sometimes (a few times a week)
    - Rarely (less than a few times a month)
    - Never/don't have

- How often do you use SMS text messaging?
    - Everyday
    - Sometimes (a few times a week)
    - Rarely (less than a few times a month)
    - Never/don't have

- Have you received any education or training on identifying phishing attacks?
    - Yes
    - No

[if they answered yes]:

- Where did you receive this training?
    - At the workplace
    - Through a formal course
    - Looked it up in free time out of my own interest
- When did you most recently receive this training?
    - Within past 6 months
    - Last 5 years
    - 5+ years ago

Have you ever been a victim of a social engineering (phishing) attack?

- Yes
- No
- Unsure

Participant Evaluations of Social Engineering Attacks:

The following are screenshots of real or fabricated social engineering attacks. For your safety, please do not enter any of the URLs you see in this survey. Additionally, we will we ask you to interact with these examples.

Please keep in mind that we are not testing your knowledge about social engineering (phishing) attacks, but want to see how you understand them.

*These questions will appear for every attack screenshot:

- Have you ever seen a phishing attack like this before?
  - Yes
  - No

- What would you do if you encountered an attack like this?
  - Delete or ignore it
  - Click the links/call the number to investigate it
  - Verify to see if it's legitimate
  - Block or report it
  - Change your password to this account
  - How confident are you that your actions will prevent you from potential consequences? (Rate from 1 – not very confident to 5 – very confident)

- Who is the most likely recipient of this attack?
  - Anyone
  - Specific group based on shared interests, demographics, etc.
  - Just the one recipient
- How confident do you feel about your response?
  - (Rate from 1 – not very confident to 5 – very confident)

- What technique is being used to persuade the reader to believe that this attack is legitimate? (select one)
  - Fear
  - Greed
  - Pretexting (i.e., giving a fake scenario to the target)
  - Curiosity
- How confident do you feel about your response?
  - (Rate from 1 – not very confident to 5 – very confident)

- At what point would you be in danger if you interacted with this item?
  - When you receive the item
  - When you click the link
  - When you enter personal information

    - When you download an item
- How confident do you feel about your response?
    - (Rate from 1 – not very confident to 5 – very confident)

- What is this attack trying to gain from victims?
    - Money
    - Personal information
    - Hack into an account/device
- How confident do you feel about your response?
    - (Rate from 1 – not very confident to 5 – very confident)

- Do you think this item is generalized (i.e. sent to many individuals, regardless of interests), or targeted (i.e. just to the recipient or based on their interests)
    - Generalized
    - Targeted
- How confident do you feel about your response?
    - (Rate from 1 – not very confident to 5 – very confident)

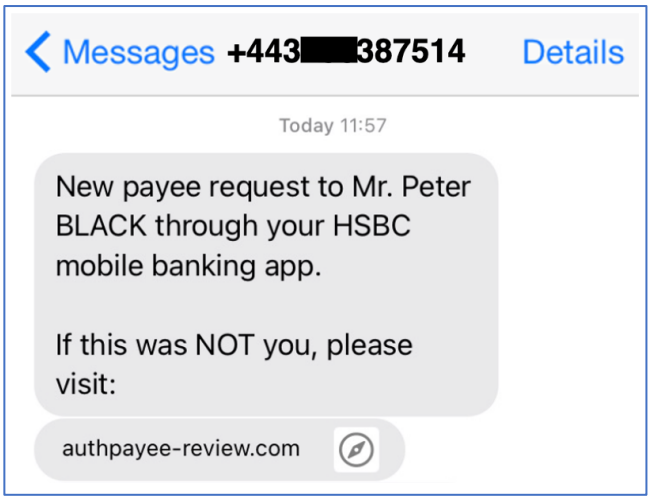
- If you saw this in real life, how likely would you be to believe it was legitimate?
- How confident do you feel about your response?
    - (Rate from 1 – not very confident to 5 – very confident)

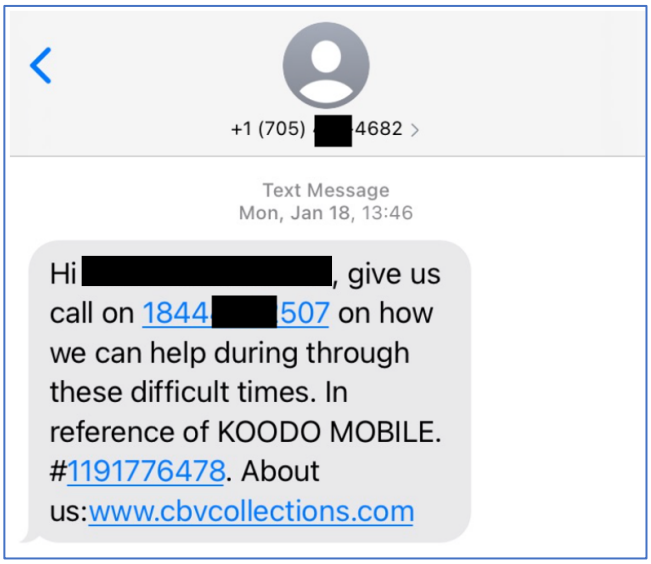
- What makes you think this example is legitimate or not? [select all which apply]
    - Spelling and grammar
    - Logo, design is unprofessional/professional
    - Offer seems unrealistic/realistic
    - Sender seems legitimate/illegitimate
    - The presence/absence of a link or call to action
    - Other: please specify

- How vulnerable do you think others are to falling for this kind of attack?
    - (1 – not very vulnerable to 5 –

very vulnerable)

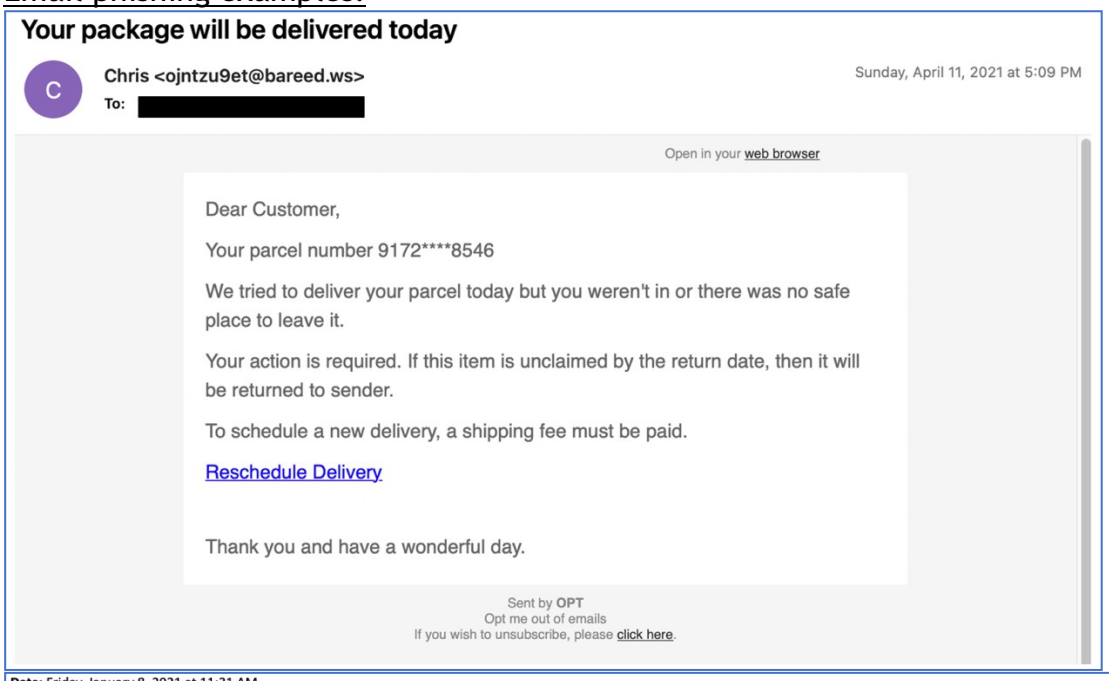
SMS phishing examples:






Lindsey
online
Hii mum xx 2:40 PM
My other phone has suffered water damage and this is my new number 😢
You can delete the old one 2:41 PM
Are you at home? 3:01 PM
Yea 3:24 PM
Need you help, got a problem 😥
3:24 PM
Whats up 3:25 PM
Can you do a payment for me 3:25 PM
Why whats happened? Ring me
3:26 PM
Now worries it's just for 2 days. For an old loan. Cant ring now. Will ring later
3:28 PM
It ain't my info but of the payment:
Name: P. Thakkar
Sort code: 04-00-04
Account no: 962[redacted]3
Amount: $2,467.00 3:31 PM
Can you message me when its done
3:31 PM

Email phishing examples:



Your package will be delivered today

C **Chris <ojntzu9et@bareed.ws>** Sunday, April 11, 2021 at 5:09 PM

To: [redacted]

Open in your web browser

Dear Customer,

Your parcel number 9172****8546

We tried to deliver your parcel today but you weren't in or there was no safe place to leave it.

Your action is required. If this item is unclaimed by the return date, then it will be returned to sender.

To schedule a new delivery, a shipping fee must be paid.

Reschedule Delivery

Thank you and have a wonderful day.

Sent by **OPT**
Opt me out of emails
If you wish to unsubscribe, please click here.

Date: Friday, January 8, 2021 at 11:31 AM
To: [redacted] <[redacted]>
Subject: You have been selected to get an exclusive $90 offer!

LIMITED TIME OPPORTUNITY

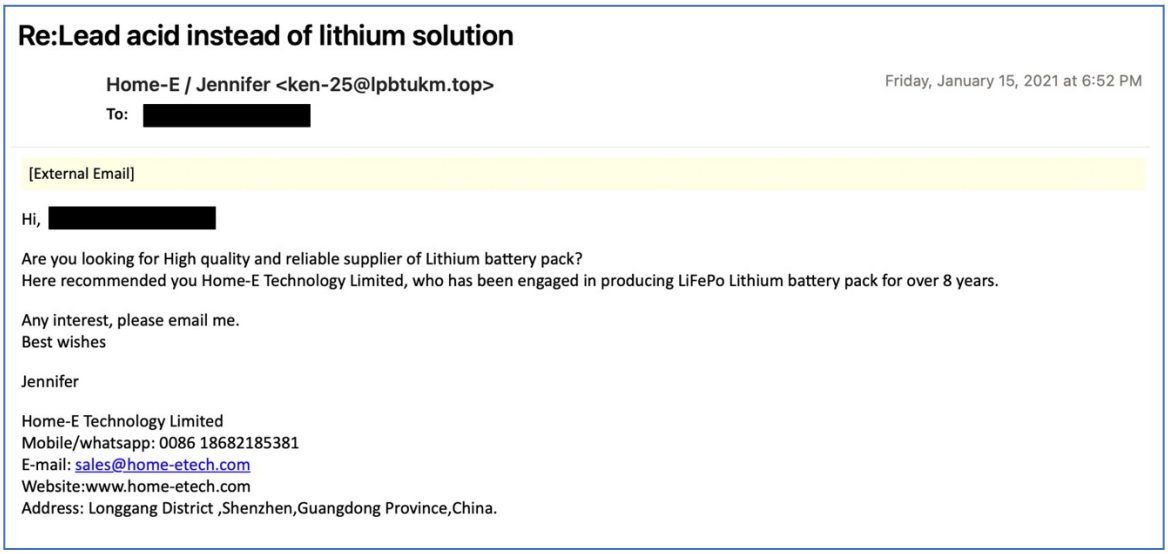
Canada Postal User!

You have been selected to get an exclusive $90 offer!

To qualify for this special offer, simply complete our 30-second marketing survey about your mailing experience.

START NOW

If you no longer wish to receive these emails, you may unsubscribe by clicking hereor by writing to 801 US Highway 1, N Palm Beach, FL 33408.



Re:Lead acid instead of lithium solution

Home-E / Jennifer <ken-25@lpbtukm.top> Friday, January 15, 2021 at 6:52 PM

To: [redacted]

[External Email]

Hi, [redacted]

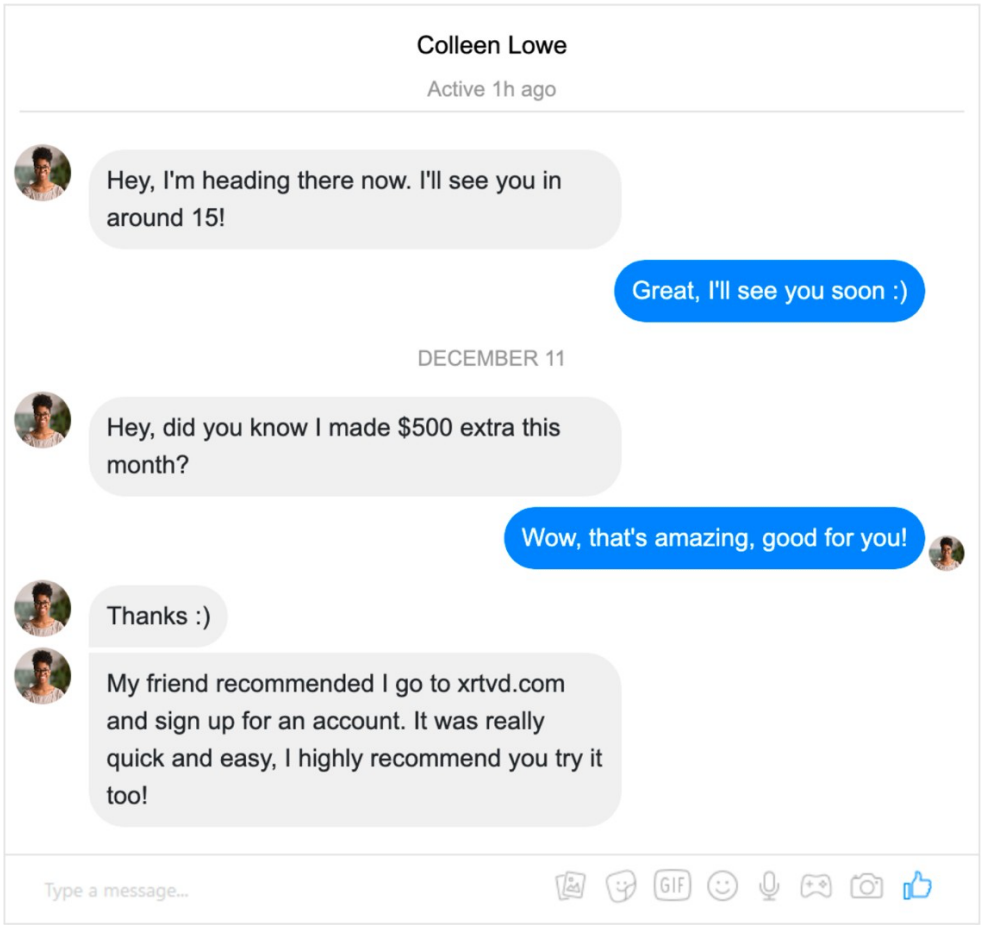
Are you looking for High quality and reliable supplier of Lithium battery pack?
Here recommended you Home-E Technology Limited, who has been engaged in producing LiFePo Lithium battery pack for over 8 years.

Any interest, please email me.
Best wishes

Jennifer

Home-E Technology Limited
Mobile/whatsapp: 0086 18682185381
E-mail: sales@home-etech.com
Website:www.home-etech.com
Address: Longgang District ,Shenzhen,Guangdong Province,China.

Social media phishing examples:



UNICEF
March 22, 2021 ·

We are happy to announce that we will be offering an ongoing internship program available to students and staff of colleges in the USA and CANADA. It's a work from home internship program and pay is $350 weekly. You can get more details and apply below:

Apply here: UNICEF.com/internship2021

Nous avons récemment reçu une notification de l'UNICEF concernant un programme de stages en cours disponible pour les étudiants et le personnel des collèges aux États-Unis et au CANADA. C'est un programme de stages à domicile et le salaire hebdomadaire est de 350 $. Vous pouvez obtenir plus de détails et postuler ci-dessous:

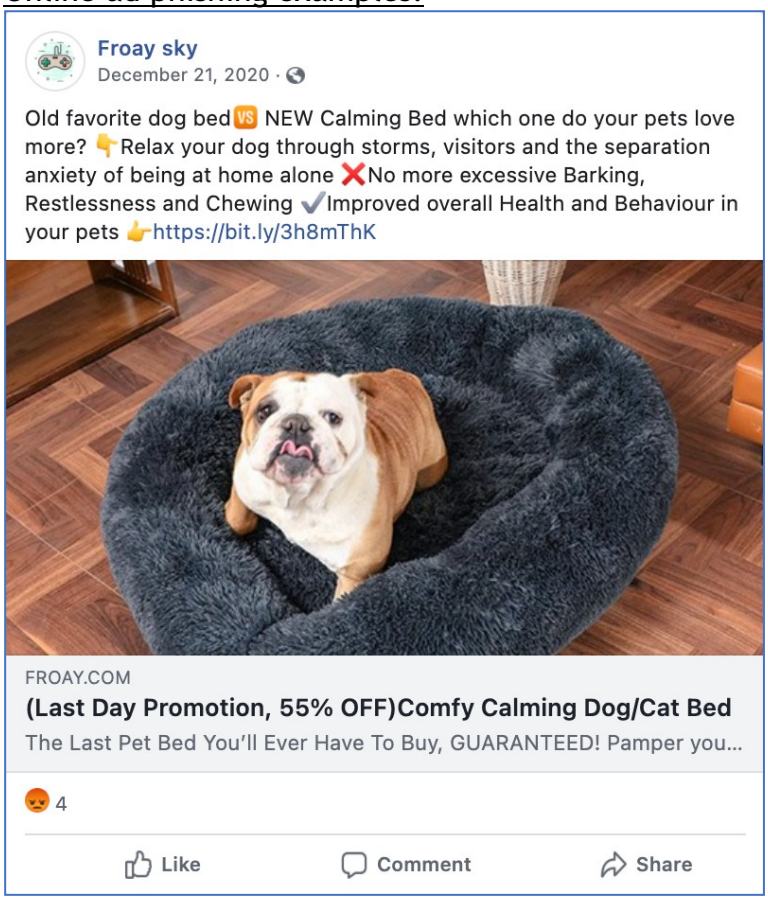
Apply here: UNICEF.com/internship2021

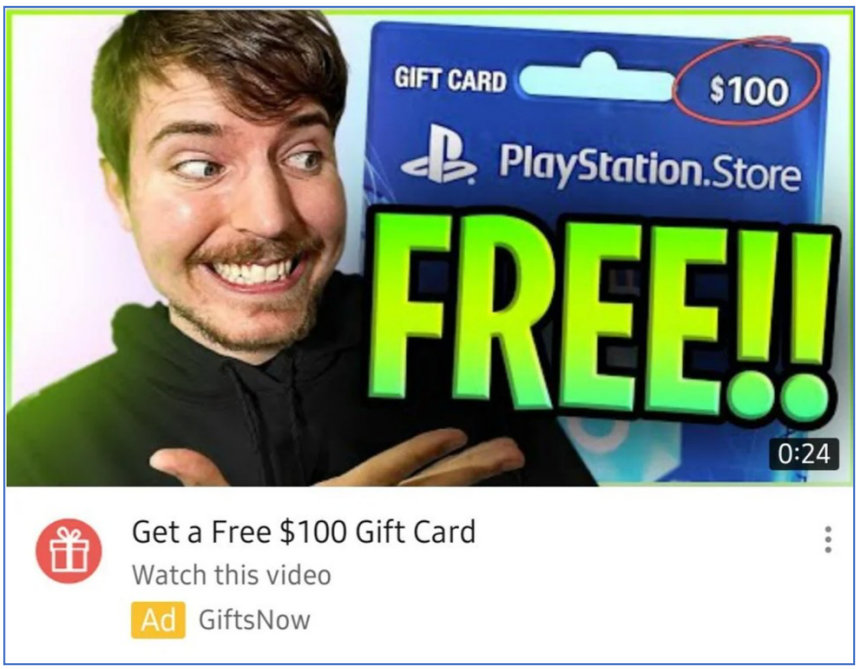
Like · Comment · Share

58 people like this.

31 shares

Online ad phishing examples:



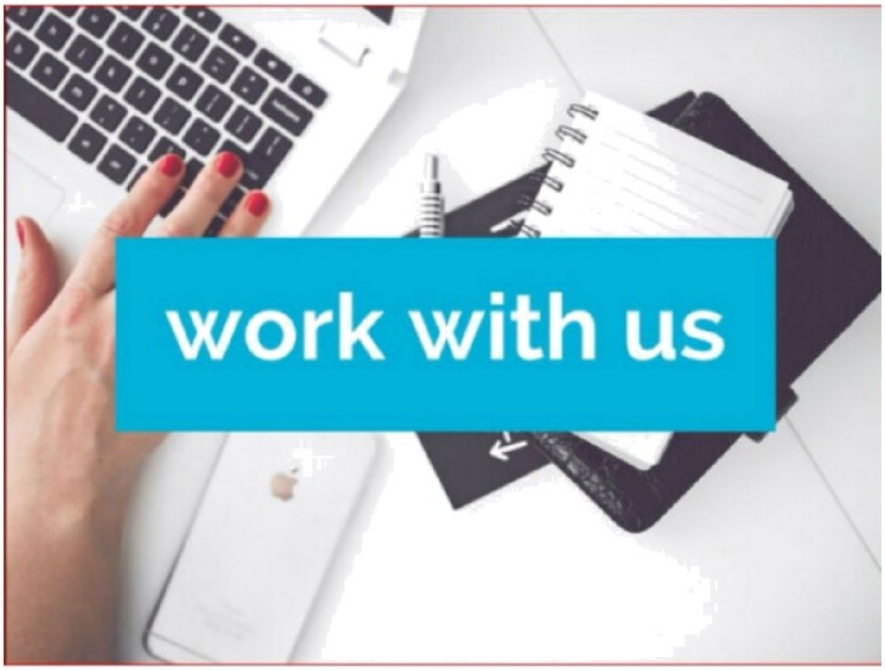


kijiji

Ontario › Toronto (GTA) › City of Toronto › Jobs › Construction & Trades in City of Toronto › Ad ID 1563622700

LOOKING FOR SKILLED TRADES!

Posted 6 minutes ago
Toronto, On



K Kijiji User

Individual

1 listing

(647) 371-XXXX Reveal

-- avg reply | -- reply rate | 1 yr on Kijiji

Job Type: Please Contact

Description

We are a general contractor looking for sub-trades to do our jobs
** REGISTER NOW AND GET SOLD JOBS ON YOUR PHONE - No leads, no nonsense. JOBS just sold ** ACCEPT OR REJECT **
We are moving to a digital platform to make it easier for our industries to accept jobs. Each order is sold, backed up and displayed on your phone as a notification that you can accept or decline
Download our pro app and register. You received sold orders on your phone today.
No more marketing and sales, payment is safe and guaranteed.

iOS:
https://cutt.ly/0xHfnK1

Android:
https://cutt.ly/IxHfJtX