Engineering and Technology Journal e-ISSN: 2456-3358

Volume 07 Issue 06 June-2022, Page No.-1270-1281

DOI: 10.47191/etj/v7i6.03, I.F. – 6.484

© 2022, ETJ

A Survey of Social Engineers that did not Collect Enough Data

Adrian Dana Austin[1], Tijjani Mohammed[2], Te-Shun Chou[3], Carolyn Dunn[4], Michael Behm[5]

[1,2,3,4,5] Department of Technology Systems, East Carolina University, Greenville, North Carolina, USA

ABSTRACT: In The context of Information technology and Cyber Security a nexus of computers and humans occurs as Social engineering. Defenders would like to know how this group interacts with users to gain access to secure information. The purpose of this research initially was to find out from Social Engineers what were the best defensive techniques. A survey was created to ask social engineers about specific defensive techniques that are prevalent in Cyber security. Not enough data was collected from social engineers and the second half of the paper covers why this was. One of the main reasons for lack of data was due to a lack of in-group trust of the researchers by the group being researched. Much of the theory behind the why has been applied to other groups in other fields.

KEYWORDS: Cyber Security**,** Social Engineering

1. INTRODUCTION

Cyber Security is a subset of the field of Information Technology. Within Cyber Security there is a subset that falls at a nexus of Human and computer interaction. This nexus is at a point where deviant actors will use social techniques to gather information to perform some deviant act. This is called Social Engineering. This subcategory includes phishing in all its forms, baiting and impersonation [1], [2], [3], [4]. Defense against social engineering typically falls into one of three main fields; policy, information technology tools, or user education [5], [6].

Education and correction of activities undertaken by humans that negatively affects the security of information technology systems is the area of study for social engineering defense. Humans are both the greatest asset and the biggest problem with information security. The breadth and depth of knowledge about this issue varies greatly. While user education is the most widely discussed portion of a defense strategy against social engineering, policies and tech tools are also important. Given the state of social engineering attacks and their rise, more data needs to be gleaned about how persons performing social engineering attacks feels defensive techniques are working. Much of the existing research has defined what techniques are being used currently, but not so much about their effectiveness.

So here lets begin by defining what phishing, baiting, and impersonation are. First phishing is best defined by [7]

The Lure. This first component consists of a phisher spamming a large number of users with an e-mail message that typically looks convincingly to be from some legitimate institution that has a presence on the Internet. The message often uses a convincing story to encourage the user to follow a URL hyperlink encoded in the e-mail to a website controlled by the phisher and entices the victim to provide it with certain requested information. The social engineering aspect of the attack normally makes itself known in the lure, as the spam gives some legitimate sounding reason for the user to supply confidential information to the website that is hyperlinked by the spam.

The Hook. This typically consists of a website that mimics the appearance and feel of that of a legitimate target institution. In particular, the site is designed to be as indistinguishable from the target's site as possible. The purpose of the hook is for victims to be directed to it via the lure portion of the attack and for the victims to disclose confidential information to the site. Examples of the type of confidential information that is often harvested include: usernames, passwords, social security numbers in the U.S. (or other national identification numbers in other parts of the world), billing addresses, checking account numbers, and credit card numbers. The hook website is generally designed both to convince the victim of its legitimacy and to encourage the victim to provide confidential information to it, with as little suspicion on the victim's part as possible.

The Catch. The third portion of the phishing attack is sometimes known as the kill. It involves the phisher or a catcher making use of the collected information for some nefarious purpose such as fraud or identity theft. [7, pp. 5-6]

For clarity vishing, is phishing done through the use of a phone conversation, v for voice. Smishing is phishing done through sms text messaging (the text messaging found on a smart phone). Baiting is where the bad actor uses a physical media and relies on the curiosity or greed of the victim [8]. [8] [4] further stat that this typically involves a USB drive specifically left in an area to be found by a victim. The media is then plugged into a system, after which the malicious software starts up and infects the system and pivots to other systems in the network. Impersonation includes almost all the other options where a person talks face-to-face with the person being engineered. This category implies that technology was used for purposes other than contact. In these cases, technology may be used as props to gain access. It should be noted that a subcategory called tailgating used by [8]would fall under this section. Tailgating includes the idea that a person acts as if they belong and should be allowed access to the same areas as the person they are following.

Social engineers are persons that perform a wide variety of tasks that relate to human interaction. For this research, we limited the scope to activities related to acquiring data and/or access to secured information systems. Social engineers also make the time to study the individual or organization being targeted before making initial contact. They may be studying the entities trash, social media information, or other publicly available information. They may also be performing varying forms of surveillance. These individuals take the time to study verbal and nonverbal cues when interacting with other humans. Interactions are not always in person, as an interaction could occur through e-mail, SMS text, voice applications, or with a software program that a person created without paying attention to security needs.

Social Engineering attacks occur frequently and happen all over the globe. These attacks can be started with research to breach local installations or with the click of a button to breach some remote location where personnel are not even aware of the attack. These social engineering attacks can occur at any time, day or night, and can even occur when the target is asleep.

The current state of defensive techniques against social engineering is largely insufficient. Hackers and social engineers are trying to gain information from the humans involved in the security system. This is because the technology and systems have been hardened to the point where humans are the weak link in the system. There have been many reviews of current defensive techniques and how they go about solving the issue of social engineering breaches. The question we asked within this research is: According to social engineers, what is the best defensive technique available against social engineering attacks?

To give statistics to this discussion, according to [9]

> [I]n 2021 Cybercriminals used social engineering in 98% of attacks... 75% of companies world wide were victim of phishing attacks in 2020… A ransomware attack is successful every 11 seconds. 60% of employees in the US click on emails even if they think them suspicious… Around 17,700 is lost every minute due to phishing. That equates to 1,062,00 million per hour and 25,488,000 million per day...The US government allocated nearly 19 billion for cybersecurity in 2021.

The need to reduce social engineering attacks is not about reducing the attempts bu reducing their success rates. Who better to ask than the people who are succeeding? The purpose of asking the validity of current defensive techniques is to ask a follow-up question to the social engineers about what they feel would be a viable solution to social engineering attacks.

The desired future state to be discussed with this research is to continue a conversation that began with what current defensive techniques are and if these techniques are effective. If a discussion between social engineers and researchers can happen, that may decrease the use of social engineering as a point of entry into information technology systems. A Second thought here is that Much of Computer science has been covered from the 1's and 0's side, or from the purely psychological side. The meshing of the two has been left out in the cold so to speak. A conversation that begins a mesh of the two sides would help give a whole picture approach to the study of the human and computer aspects of computer science. Meshing of the two sides might include such ideas as why certain libraries (pieces of reused code) are always reused and not tested for malicious activity on a regular basis (this is beginning to be done). Also why are there the same kinds of holes in code from new programs as there were found in older codes, did a new generation of coders not learn all the lessons of the previous generations? And errors in hardware design causing external access to a computer simply because the right testing may not have been done in a rush to market. But this paper focuses on social engineering, one aspect of the human side of computer science.

2. SURVEY METHODOLOGY

To begin with our survey, it was found no one that the researchers could find had asked social engineers their perspective on defensive techniques. This idea of asking the deviant persons about their deviant behavior has been used in other areas. From here the idea was to do a pilot survey to get a beginning down on paper of where research on the specific topic should go. A mixed methods approach was chosen to allow for qualitative and quantitative results, giving the broadest overall picture of the desired results.

2.1. The Survey

A nominal variable question was chosen to start off the survey:

1. What forms of information gathering do you perform?

a. Online search using OSINT (Open Source INTelegence)
b. Physical Gathering of information through surveillance or going through trash or other physical means
c. Both A and B

This question was chosen to be first as a test question, as someone who is a social engineer would typically answer either A or C and someone who is a confidence man would typically answer B and therefore not be a social engineer, typically.

Next was a series of Likert scale questions about multifactor authentication, e-mail filters, system patching, good physical security, good security policy, limiting access by unknown hardware, and user education. These questions were about whether the above were positive defensive techniques as part of a multi-faceted approach to security. The quantitative portion of the question was followed up with a qualitative portion asking if the survey taker found a specific technique in these categories that works better than others.

A brief discussion of the techniques mentioned in the Likert scale questions is as follows.

Multifactor authentication is a method of providing one's identity that requires more than just a password. "To palliate password weakness, multi-factor authentication protocols combine several authentication factors. Typically, instead of using a login and password, the user proves possession of an additional device, such as his mobile phone, or a dedicated authentication token" [10]

E-mail filtering is a way to keep out many phishing attempts from even reaching the end user. [11]break down e-mail filtering into three categories: basic features, latent topic model features, and dynamic Markov chain features. Each of these categories is a way to keep social engineers from phishing end users but is beyond the scope of this research project.

For system patching [5] states is best by saying: [M]issing security patches are one of the biggest problems that allow successful exploitation … to give yourself or your computers the best protection against software vulnerability exploitation, all you have to do is apply security patches in a timely and consistent manner … Unfortunately, effective patching remains overly difficult and elusive.

[5]also states the following about patching

1. Most exploits are caused by old vulnerabilities for which patches exist.
2. Most exploits are caused by a few unpatched programs
3. The most unpatched program is not always the most exploited program
4. You need to patch hardware too.

[5]Also lists some common patching problems:

1. Detecting missing patching isn't accurate.
2. You can't always patch.
3. Some percentage of patching always fails.
4. Patching will cause operational issues.
5. A patch is a globally broadcasted exploit announcement.

Physical security is something just as important as cyber security of your networks. [2], [3], [4], [6] all discuss how to get around physical security in their respective books. A majority of the concepts behind social engineering can be used either through electronic means or in person. Having good security training for your people that provide security and having appropriate locks and security systems are equally important.

[5], [6] discuss security policies and contraols as defensive measures. A good security policy includes "Clear instructions that provide the guidelines for employee behavior…and are a fundamental building block in developing effective controls to counter potential security threats." Polices and controls should be written so that non-technical people can understand them. [5], [6] discuss how explaining policies and breaking them down for non-technical people will make them more likely to be followed. Understanding why (explained appropriately) will keep users from just bypassing polices they feel are just broken.

Limiting unknown hardware is a defense that speaks directly to the baiting technique, mentioned at the beginning deals with this limiting. Here reducing the ability to connect any hardware to the network, or your computer that has not already been verified is the defensive technique. Unknown hardware could be any electronic equipment that can attach to your computer or network such as USB drives, CD drives, audio players, smartphones, tablets, laptops and others.

User training is best described by [6]as follows: Security training must have a significantly greater aim than simply imparting rules. The training program designer must recognize the strong temptation on the part of employees, under pressure of getting their job done, to overlook or ignore their security responsibilities. Knowledge about the tactics of social engineering and how to defend against the attacks is important, but it will only be of value if the training is designed to focus heavily on motivating employees to use the knowledge. The company can count the program as meeting its bottom-line goal if everyone completing the training is thoroughly convinced and motivated by one basic notion: that information security is part of his or her job.

All of these questions were followed up by qualitative sections that allowed for the survey taker to describe a particular technique in each of these subsections. Asking if the survey taker had found a specific technique that they felt works better than others. This section transitioned into some open-ended qualitative questions about social engineering techniques and the best defense against them. Just to be clear we are distinguishing between defensive techniques and defense against offensive social engineering techniques. These offensive techniques, which are communication modeling, pretexting, influence, manipulation, elicitation, nonverbal cues, and reverse social engineering.

Communication modeling is a technique discussed by [2], [3]. Communication modeling is about how to interact with other people. [2], [3] breaks his model down into three parts. The first part is the approach which involves knowing how to present oneself to someone with whom one wishes to interact. "It is those first crucial seconds of interaction between you and a stranger that will set the tone for the rest of the engagement." [3] Sizing a person up and determining what can be ascertained based on their attributes and then deciding upon the approach to get a positive outcome can be very difficult. In the approach, one needs to be able to answer the following four questions in the first 5-10 seconds of the interaction: "Who are you? What do you want? Are you a threat? How long will this take?" [3] Secondly, [3] utilizes his DISC acronym using four quadrants of a circle as Direct/Dominant, Influencing, Supporter/Steady, and Conscientious/Compliant. The wording of the acronym quickly demonstrates the different types of communicators involved. Thirdly, understanding where the mark or victim lies within the circle and where the social engineer lies within the circle can also influence the outcome of an interaction between the two. The social engineer will be more successful and gain a positive outcome if they use the first few moments of an interaction to determine what quadrant of the circle the person being engineered falls into and how that will mesh with the engineer's own location on the DISC circle. There are techniques [3] says to used based on where the two person fall within the circle, but that is outside the scope of this document. Finally [3] discusses the limitations of using his DISC acronym. He discusses how this is not some sort of magic wand and that communication modeling takes time. The good news, according to [3], is that this will work in all of the areas of social engineering mentioned at the beginning of the paper (i.e. phishing, baiting, impersonation)

In his book, [4] devotes an entire chapter to the topic of pretexting. He states that research and planning go into performing this particular activity. In addition, he discusses body language and non-verbal cues that influence its effectiveness. Pretexting involves creating a scenario that is presented in a manner that is believable and induces trust by the persons being engineered. This scenario includes the social engineer's background information and potentially using disguises and identity impersonation. [4] also states that pretexting goes beyond simple flattery or being ignorant of a topic or subject.

Manipulation is a form of influence. In an article published by Forbes, [12] interviewed Bob burg, a famous author, on influence and manipulation related to business. He stated that the actual divide is between manipulation on the negative side and persuasion on the positive side. In the article, Burg states that influence in and of itself is neutral.

> It's sort of like the physical law of gravity. Gravity in and of itself is neutral. It manifests itself as good when keeping us floating aimlessly up into space. It manifests itself as bad when we fall off a seven-story building… Both manipulators and persuaders understand human nature, human motivation, what drives people to take action on certain ideas. But while manipulators will utilize that knowledge for their own ends only, persuaders will never do that. [12, para 14]

[13], [4], [3] go into a bit more depth on the following subtopics in their work: authority, concession, likability, obligation, reciprocity, and scarcity. Regarding authority, social engineers position themselves in a position of authority over the intended target in some manner. This form of social engineering typically works in larger business entities and when the social engineer is not physically present. In an influence operation that includes concession, [13]defined it as letting of of something you appear to want and settling for something smaller-in other words, telling the subject being social engineered that you want something big and a little out of reach and when they say no, accepting that and asking for something within reach and smaller. This smaller thing that is withine reach is actually what the social engineer wanted in the first place. Likability is an influence technique that involves the engineered feeling liked in some manner. With likeability, it can be difficult to judge the level of compliment necessary to keep the person from being too skeptical. [4] discusses how following up a small compliment with a question is a good example of gauging the proper level. The example he uses is "That's a great looking watch. May I ask where you bought it?" Obligation types of influence including having the subject being engineered to feel obligated to respond to questions that are asked due to, in some situations, societal norms. These societal norms include gratitude or feeling that they owe the engineer the information. According to [4], this could be something as simple as holding the door open for the target. Obligation also includes the idea that the social engineer may give up tidbits of information that appear to be about themselves, thereby obligating the one being engineered to reciprocate. With reciprocity, one must simply apply the rule of treating others as you would wish to be treated. The previous example of giving up information about oneself could also fall under reciprocity. "This feeling of indebtedness triggers reciprocity in your target and makes them much more likely to fulfill a request" [13, para 5]Scarcity is the last concept on which the three authors agree. [4] succinctly states:

In social engineering, scarcity is used to create a situation or feelings of urgency necessitating the target to make a quick and rash decision. Of course, the scarcity situation itself is one that is fabricated by the social engineering and the choices provided are not in the best interest of the target. The desired outcome is one that forces the target to go against their instinct and comply with the social engineer's request.

The two additional topics under influence that only [13], [2] discuss- social proof and commitment & consistency. Simply put, social proof states that everybody is doing it, and you should as well. Social proof is the idea that a target will perform a task or provide information based on the idea that everyone else is doing said task or providing said information, and the target does not with to be left out-in other words, using the concept of herd mentality against a target for fraudulent gains of some kind. The last subtopic mentioned by [13] [3] is consistency & commitment. In this subtopic, the social engineer knows that the target wants to be consistent with their answers to the social engineer. For example, the social engineer with start with something small, with which the target will comply and build from there to the information they truly wish to know. Starting small and inconspicuous will get the target going in the right direction. The target will wish to be consistent with their interaction, allowing the social engineer to gather the information they are looking for in the conversation.

Elicitation is the art of getting information without asking direct questions. Both [2] [4]define this in similar ways; it is about having a regular conversation discussing typical topics and throwing in leading questions that allow the victim to offer the information the social engineer is seeking. According to [2]there are ten techniques he refers to from a book by [14]). [4] advises many of the same techniques, and the ten are: Artificial time constraints, Accommodating nonverbal cues, slower rate of speech, sympathy or assistance, ego suspension, validation, quid pro quo, reciprocal altruism, managed expectations.

When talking to a subject, the engineer will create artificial time constraints, where they will appear to have someplace else they need to be in a short time. This will make the subject more comfortable because they can clearly see that there will be an end to the conversation. This clear ending point makes it so that the subject can feel more in control and able to realize that with a few short conversing words, they can quickly get on to what they were doing previously. [14]states that it is about threat level instead of controlling the situation. [14] [4] [2]) all talk about nonverbal communication, specifically body language. The idea here is that you appear approachable in your stance and attitude. There will be more discussion about body language later in the chapter. Another nonverbal cue that is discussed is a slower rate of speech. [14]states that speech can be changed and that the speed at which we speak can affect how listeners will view the speaker. He uses the analogy of fast-talking to that of a used car salesman. More rapid speech is regularly associated with someone attempting to sell something not as high in quality as they would have the buyer believe. Therefore, a person with slower and more deliberate speech is perceived as honest and forthright.

With the idea of sympathy or assistance in elicitation, the thought is to find [a] third-party reference … a topic used to initiate that isn't too personal about the individual targeted for discussion. The topic is also not about you. Individuals typically do not like talking to strangers about either of these topics, at least not in the first few seconds. [14, p. 36]

Ego suspension is the concept of suspending your egotistical thoughts and putting the wants and needs of another person ahead of your own in the interaction. This can deescalate the situation where a conversation may occur when otherwise it would not have occurred. Additionally, putting the engineer's ego on hold can elevate the other person's ego whereby they may continue to give information, after which otherwise they would have ceased. [14]splits validation into three components: listening, thoughtfulness, and validating thoughts and opinions. The first is simply the act of listening. This listening is a way to validate what the person is saying in a simple form. The next is thoughtfulness, which includes providing small gestures that show one is placing another's needs, wants, and welfare above their own. [14]states this in a simple example of having hand sanitizer or chewing gum and offering them a portion of the items during a conversation. The final component of validation is to validate the thoughts and opinions of the target. [14]provides this example:

While at the meeting, I asked my source, "So, what do you think about country X?" His response was perfect. He said, "I think they are doing great harm to the United States." I responded, "That's an interesting point of view, why do you think that?" Following his response, I validated his thoughts again, and then asked him what he thought we could do about it. The entire dialogue and process was centered on my source's ideas and me validating them to have him take action. [14, p. 54]

Specifically, ask how, when, and why questions because there is a socially accepted way to ask these questions. Once you have led with the other techniques to get the conversation started, the how, when, and why questions keep the conversation going. For clarity, these are not one-word questions you should ask but the beginning of a longer question that shows interest in the topic of conversation.

Alternatively, the question goes from the path of the first conversation down a path that veers away from the original path towards another topic of conversation. This concept of quid pro quo is the Latin phrase "something for something." The idea here that all three [14] [2] [4]bring up is that when the social engineer is interacting with the subject, they must give up some information about themselves to get the subject to also offer information. The social engineer should not push too hard with this but to "slowly build trust through non-threatening dialogue" [14, p. 67]. With the technique of reciprocal altruism, also known as gift-giving, the social engineer will give the subject something. Reciprocation of that gift is a psychological need. An example is when someone holds open the door. A social person would, in turn, hold the door open for the first person at some nearby time, such as when there are double sets of doors.

This final technique [Manage Expectations]is one where being able to mask one's actual agenda or shift the agenda to appear to be altruistic is a positive for the social engineer. Every conversation or engagement with another human being has an agenda. Another definition of agenda might be objective or desired outcome. Sometimes the agenda is to sell you a used car. Sometimes the agenda is to share a secret. Other times, it is simply to make another person feel better. Regardless of the situation, whether it is an altruistic intention or not, there is an agenda. The individuals in life that are able to either mask their agenda or shift the agenda to something altruistic will have great success at building rapport. [14, p. 77]

After elicitation and all of its parts, is Nonverbal cues, they have been touched on in elicitation, but deserve their own small section for emphasis. For a social engineer, most nonverbal cues typically happen in a face-to-face meeting. Many of these subtle expressions of emotion are lost when the discussion moves to the written word or text messages. "These nonverbal cues or nonverbal communication include facial expressions, gestures, haptics (touching), physical movements (kinesics), posture, body adornment, tone, timbre, and volume of the voice as well as previously mentioned the speed of speech" [15, pp. 2-4] The final technique was left off of the survey but for completeness sake we are including it here. This final technique is reverse social engineering, here, the engineer makes contact with the subject and implies that they can solve a problem for the subject. This could be a current problem that the subject has called them about or a problem that the social engineer creates later, or one that happens to arise.

From here the survey moved on to two nominal variable questions.

1. Which of these defensive techniques warrants more user education than others?
   a. Password creation techniques
   b. Multifactor Authentication
   c. Defense against phishing
   d. Defense through Security Policy
   e. Defense through Physical security awareness programs
   f. Other [Open ended]
2. Of the vectors that have been used to start a Social Engineering contact which do you feel is most productive
   a. Phishing
   b. Vishing
   c. SMishing
   d. Impersonation
   e. Other [open ended]

These two questions were included because in the literature user education came up as the most needed defense, and a specification of which technique needed more education seemed warranted in a pilot study. The second question was added as for a pilot in asking how does one start a social engineering technique. The literature states that these are the most used starter techniques but this is from the view point of the author in industry and not from a researched, survey point of view. The next section of the survey was open-ended questions about phishing, vishing, smishing, and impersonation. This section was included due to the previous nominal variable question to ask of the social engineering vectors asked what do you feel is a positive defensive technique. Next for the survey was the demographic questions. These questions were kept vague as to keep from eliciting a negative response to answering them. Job title, general geographic location, Gender, and age bracket were all asked in such as way that someone could answer them and be in a category that could be generic enough to not be able to correlate to a specific person, but specific enough for the survey purpose of analysis. After this section I asked if there was any thing else that I should have added to the survey. And then to incentivize people to give an email for a possible follow-up qualitative interview by some future research, I offered 4 gift cards.

The survey was then sent to a few persons in industry that chose not to respond for comment. And then the researchers also reviewed it, along with a discussion with a research coach. This review was for common errors such as leading, confusing or double-barreled questions. For clarity, a double-barreled question is "one where the question is about two things that could have opposing responses." [16] The survey was determined have some redundant questions and needed to be changed or removed from the questionnaire. This was partly done due to a belief that a shorter survey would be more likely to be answered.

The survey at this point was deemed appropriate to send out, so the survey was created in Redcap and a link was created.

Study data was collected and managed using REDCap electronic data capture tools hosted at East Carolina University. REDCap (Research Electronic Data Capture) is a secure, web-based application designed to support data capture for research studies, providing 1) an intuitive interface for validated data entry; 2) audit trails for tracking data manipulation and export procedures; 3) automated export procedures for seamless data downloads to common statistical packages; and 4) procedures for importing data from external sources. [17]

The link was them imbedded in a comment posted on Reddit, a social media site that is one of the largest.

Reddit is home to thousands of communities, endless conversation, and authentic human connection. Whether you're into breaking news, sports, TV fan theories, or a never-ending stream of the internet's cutest animals, there's a community on Reddit for you. [18]

The survey was allowed to run online for 2 months, and was in the subreddits (r/Cybersecurity, r/Socialengineering) at the end it was also sent out to 214 profiles on Facebook to see if more responses could be acquired.

3. SURVEY RESULTS

There were more sociological findings than computer science findings for the surveys that were answered. However, while there were only 12 people who answered the survey (and those mostly partial answers), a direct synopsis of what was collected can be found in Appendix E-Synopsis of results. And a discussion on the results is here.

3.1. Likert scale questions.

In 1932, Rensis Likert published a paper titled "A Technique for the Measurement of Attitudes," in which he discussed a psychological measurement scale that employs questionnaires. Almost since the beginning of the usage of his scale, there have been several discussions on its use. For example, in [19], the authors discuss how there should be multiple questions on a single topic to make the validity of the research more prevalent and accurate. They also state that "single items are appropriate when the referenced concept is singular, concrete, and understandable to the respondent" [19, p. 134]. Willits et al. also discuss how a scale of "at least four are needed for evaluation of internal consistency" [20]. Moreover, while reliability measures increase as the number of items increases above five, each addition has progressively less impact on the scale reliability ([21] [22].

For the Likert scale questions, it was found that, for the most part, people marked *strongly agree* or *agree* for this section. There were two partial surveys indicated *disagree* with all of the Likert scale questions. There is no scientific evidence to support the idea that these two survey takers were trying to skew the results, however there is also no data to suggest this idea is wrong. This was the most complete section of the 12 surveys. Unfortunately, no one answered the qualitative component about which techniques work better. Even with the interviewee from NIST, they stated that they did not feel comfortable giving out their preferences. This finding is partly due to trust issues and the idea that the interviewee was not sure of how much I would reveal and who would be reading this to what end. Some of the reasons that the Likert scale was the most responded to part of the survey may be the ease of answering compared to the open-ended questions. Discussed later will be the point that technical word use may also have negatively impacted the comparative levels of the answerability of the Likert questions.

3.2 Open-ended questions.

Open-ended questions in this survey were used to gather further information. Although these questions were included to help elucidate answers to quantitative questions,

> The literature on open-ended questions has established that answering these questions places a greater burden on respondents' cognitive abilities than selecting a response category in a closed-ended questions, since respondents must formulate the answer in their own words and express it verbally or in writing. [23, p. 466]

Additionally, they found no significant difference in the level of response to open-ended questions based on the number of questions. They did state that this may be affected by being able "to recruit participants who are trained in answering survey questions (and cognitive probes)" [23, p. 466] Other research suggests how many questions are answered is based on the quantity presented. The open-ended portion of the Likert scale questions were not answered. Question 9 (What do you feel is the best defensive technique against Communication Modeling and why?) of the opened ended only section was the only one answered, and by only one person. Their answer:

> User education in combination with strong IT sec[urity] practices. Getting people to retain some amount of skepticism towards what's asked of them in combo with IT making sure standards are followed so that those in authority are never asking for PII over unapproved channels.

This person also asked how this question on Communication Modeling was different from the others discussing Pretexting, Convincing, Influence, Manipulation, Elicitation, and Nonverbal cues. This tells the researcher that what the person stated was important for them to state, however, the response gives proof that they may not have understood the difference between the techniques. Based on research into the sociology of why people do not answer these questions, and the interview with the person from NIST, these questions were most likely not answered because of the lack of knowledge of how the questions were going to be used.

For the section of the questionnaire involving what works for defense against Phishing, Vishing, and Smishing, the one person to answer put the same answer for all of them: "User education in combo w/ consistently followed protocols within the company." This consistent thought gives rise to the idea that the research may have been too granular for these as phishing, vishing, and smishing are all basically the same thing using different platforms. This section shows that this person sees user education, security protocols, and consistent implementation of security protocols as the best defense to phishing, vishing, smishing.

For the last question in this section about Impersonation this same person stated:

> Giving people an easy way to look up a person within their organization. Impersonation works best in large organizations where it's conceivable to interact with someone you've never met before who has just been hired yesterday. If you're worried about this kind of attack, making sure people have swift access to an org chart is essential [.]

This is a well-thought-out response and it tells the researcher that Impersonation is something this person sees as a high potential of happening. This answer gives more credence to the idea that phishing, smishing, and vishing should have been one question and that this could have shortened the survey.

Of the 12 surveys, only one person answered this section. It shows that to that one person this was important enough to take the time to answer. However, there is not enough data to give statistical evidence of the significance of any outcomes.

3.4 Nominal variable questions.

For our survey, we asked three nominal choice questions. The first question:

1. What forms of information gathering do you perform?

This solicited all three answers, a majority of the respondents stated both A & B. This question was not in the list of research questions to be answered but was merely an ice breaker for the survey. It was believed that actual social engineers would answer *both* or *online*. But it was shown that all three answers seemed viable. Based on my research none of the social engineers should have answered just "Physical gathering of information through surveillance or going through trash or other physical means." This seems like the answer of someone who is a confidence man, and not a social engineer because by implication there was no electronic activity used in finding out information on the subject.

Both of the following questions required specific knowledge of the topic of social engineering:

2. Which of these defensive techniques warrants more user education than others?

For this question the two persons that answered both answered that "Defense against phishing" warrants more user education. This is consistent with the fact referenced from the [24] that phishing consists of 36% of all attacks reported.

3. Of the vectors that have been used to start a social engineering contact which do you feel is most productive?

This question was answered by 2 persons. One person answered *Impersonation* and one person answered *Phishing*. Impersonation can imply in person, over the phone (Vishing), through text message (Smishing), even phishing, or more specifically spear phishing. Spear phishing, not already defined, is a form of phishing that targets specific persons for an attack.

The three questions were not answered by enough persons to warrant any form of statistical analysis. The last two questions were answered in a manner consistent with industry articles and reports, some of which were mentioned in this thesis. This consistency with industry opinion may mean that the persons that answered have read these articles or may mean that they have actually answered in earnest.

3.5 Demographic questions.

The demographic section was only answered by two people. Both respondents were female, one a public servant, one a technical program manager. One respondent was from the Triad (the piedmont region of North Carolina), the other was from San Francisco, CA. One of the respondents that answered was in the 25-34 range and the other was in the 35-44 range. Much of the reason this section was not answered may be due to the private nature of social engineering.

The section about what could have been done differently was not answered by any participants. This is partly due to individuals giving up at an earlier point in the survey. Only one person left an additional comment; a quote from a movie "be excellent to each other." [25] No other comments were made.

It should be noted that the research did not anticipate the necessity of trust relationships that are required for the completion of a successful survey with this population. The interviewee confirmed the researcher's assumption that after 2 months of the survey being available, the survey would have gotten better responses had the researcher been a known entity to the groups being surveyed. This would have been helped by going to message boards and posting, asking questions, and giving tips and hints to relevant topics to prove the researcher's in-group status.

4. DISCUSSION

In this section, the discussion covers why and how the survey did not collect enough data to reach its full potential. First, the researcher did not take the time necessary to become a visible part of the online community. Becoming a known entity in the cybersecurity community and potentially more worthy of trust would have allowed more acceptance of guidance in completing a survey. Next, potential participants may have decided that the survey was too long for the reward. Also, there are several reasons why certain questions may not have been answered; these potential reasons are discussed.

Finally, social interaction and the human aspect of computer science are side topics that have not been explored in depth. The human aspect of social engineers and their disapproval of “outsiders” is discussed in this chapter.

4.1 Social interaction

Attempting to interact with social engineers can be difficult at best. A group of people who know how to prey on the trust of others is not likely to easily assume that other people are trustworthy. This in and of itself points to a particular need to distrust others as their social engineering activities take advantage of trust for deviant activities.

4.2 Defining interpersonal trust.

The first step in understanding the human interaction for this survey is *interpersonal trust* or understanding how people trust others. This is not a simple concept to define and has been interpreted differently by many researchers. For this work, The researcher uses [26] work that states that interpersonal trust has:

> elements that referred to (1) a subject, (2) an action/behavior, and (3) a future action (i.e. an intention) and/or expectation (i.e. belief). The future element, which involves predicting or anticipating another’s actions, is a distinctive and critical feature of trust. Deception, for example, is about something that has happened or is happening. Trust, however, involves present decisions, often based on another person’s past behavior, that require anticipating some action that hasn’t yet happened. [26, p. 8]

[26] synopsis of trust research is very in-depth. From this point, he discusses several research projects that have been done in this area going back to Homans in 1958. In Homan’s project, he discusses how trust is formed by associations already made. For instance, if someone you trust trusts someone, you can begin to trust from the point of a positive trust attitude with that person instead of a zero-trust point or some negative trust beginning point. Other social beginning points for trust include predeterminations based on someone’s mode of dress and previous interactions with persons in that mode of dress. He outlines other beginning points, but they are not relevant due to the survey being online, and the only starting point for a social basis of trust would be from a screen name and first words in a chat room.

[26]also discusses how there is a neurobiological aspect of interpersonal trust. In this part, he discusses how different parts of the brain have a great deal to do with the trust spectrum. [26] explains that people who are not neurotypical do not start from a zero-trust point. That is partly due to the development, or lack thereof, of certain portions of the brain that contribute to many people being unable to start from this zero-trust point. Because of their brain development, some people are more predisposed to trust people, while people on the other end of the spectrum start from a negative-trust point. [26] also discusses how the brain, or the injection of certain chemicals, can change the point on the spectrum a person starts from given differing levels of oxytocin, vasopressin, and dopamine. While these chemicals are involved in other activities in the brain, they are also involved in trust and trustworthiness.

Here concerning [26], and due to the nature of social engineering and its innateness of negative trust, social engineers had no previous trustable actions upon which to base future actions, i.e., taking the survey. They started from a point where they do not typically trust many people outside their social group. This lack of trust is partly due to the nature of social engineering and its taking advantage of trust for deviant activities and results. Had the researcher taken the time to gain some trust in the group, there would have been previous trustable acts to base the future action of taking the survey. Another perspective is that given Homan’s ideas, had the researcher made posts to the Reddit forum in the past, joined group discussions, asked questions, and answered questions to the point of acceptance within the group, there would have been a past basis of trust. Then it would have been acceptable to complete the survey as something they were doing as a benefit for a group member to get some of the cost back later.

4.3. In-group status.

As defined in this thesis, interpersonal trust is presented to show that the NIST interviewee was correct from a scientific perspective. A person would have been more likely to answer the survey if the researcher had been associated in some manner with the group being surveyed. There would have been no need for deception, as a simple acknowledgment from the majority of the group to which the researcher belonged would suffice for In-group status. [26] cited research stating,

> Within the in-group there exists a depersonalized bond of trust that extends to all its members; one that is not contingent on other social knowledge or affective connections between individual parties. Group membership itself carries the imprimatur of trustworthiness. Some have referred to this as a form of “Category-based trust” (Kramer, 1999) and there is some evidence, as we have seen, that such a category-based trust can help reduce cognitive load in humans by providing mental shortcuts: you can trust person X because they are part of group Y. [26, p. 42]

Because of the lack of face-to-face interaction in an online chat, the people in the group would have only had this association from which to work. The ability to transfer trust is the next theory I discuss.

4.4 Trust Transfer Theory.

[27] discusses trust transfer in online contexts in her research, specifically how trust can be transferred from known individuals to unknown individuals. She also finds that it can be transferred from a place or an industry association to an individual [27]. She states,

> Campbell (1958) suggested that such perceptions are based on the similarity, proximity, and common fate of entities. He introduced the term 'entitativity' to describe the degree to which a collection of individuals is perceived as forming a group. The concept of entitativity allows for the study of collections of individuals who vary along a continuum in the extent to which they are perceived as forming a cohesive unit, rather than forcing such collections to be categorized in a dichotomous fashion as forming a group or not. [27, pp. 2-3]

This entitativity is varied in its perception from the perspective of entities within the group.

Thus, joining a group of social engineers and fulfilling a role within the group (starting as the new person trying to gather information through asking questions and eventually getting to the point where one contributes to discussion as an equal) is a goal for future research with this population. Once this point has been reached, a certain amount of trust would be transferred, leading to better survey response rates. Ideally, this would also lead to more honest and thorough survey responses for qualitative and quantitative components. To put this more aptly, the researcher sees that becoming part of the group cohesion through a shared bond would have changed the cost-benefit analysis scale in favor of group members completing the survey as a benefit of being part of the group.

4.5 Use of an Online Survey in General

The value of using an online survey is magnified for social engineers. For example, this may be the only way to contact them as they like to remain anonymous, and sending things to a street address or calling them on the phone would eliminate much of that anonymity. [28] discuss nine potential weaknesses to using online surveys, as discussed below.

4.5.1 Perception as junk mail

According to [29], 84.14% of all e-mails sent in September 2021 were spam. Because of this, survey recipients would likely believe that a survey received via e-mail is some kind of spam, specifically perceiving it as a potential phishing scheme. Part of the reason I surveyed through an interview with the person from NIST was that he stated that I was most likely going to be sending a link to the survey from a source that was not trusted enough for him to open. Because the nature of social engineers is to take advantage of others' trust in the general good nature of people, it could be inferred that social engineers would assume that the survey may have been sent by another social engineer trying to phish them.

4.5.2 Skewed attributes of the Internet population.

While Internet usage can be seen as varying greatly from high-income populations (89%) to low-income populations [30], [28] also believe from their research that Internet users are typically male. This demographic evidence indicates that the survey is more likely to get answers from affluent males than from impoverished females.

4.5.3 Questions about sample selection.

Using Reddit and Facebook as survey distribution mechanisms was inherently problematic for this population because these two platforms preclude social engineers who do not use these sites frequently enough to see the survey link. Additionally, the previously discussed issue about not being a known quantity in these online communities further skewed the sample population.

4.5.4 Respondent's lack of online experience/expertise.

In this section in [28] article, the authors discuss the potential that the possible participants would not have the expertise or experience to know that the survey was legitimate or have a frame of reference to know what taking an online survey would mean to them directly. I feel this is currently less of an issue than it may have been when the article was published. The user's expertise identifying as a social engineer would be much higher than the average Internet user. This is partly due to how the typical social engineer uses the Internet to gather information on a potential victim.

4.5.5 Technology variations.

Research by [28] has aged regarding this topic; however, the idea that technological variations would affect the survey is still a valid point. Since many, if not most social engineers want to maintain some level of anonymity online, their trust levels may vary regarding a survey creator and their ability to maintain their chosen level of anonymity. These technological variations would also alter the trust levels in the link used for the survey and may result in the need for interview-style survey responses. Interview style survey responses would most likely need to be Internet chat-based interviews; this is partly due to the need to remain anonymous by the groups to be interviewed.

4.5.6 Unclear answering instructions.

[28] discuss how unclear instructions may cause frustration by the survey taker and therefore result in the survey taker exiting the survey without completing it. This may have been part of the problem with this survey, although more likely due to the technical level of the terminology used in the instructions. While common in industry and academia, these terms may not be as widely used in the common vernacular of social engineers. Their goal is to blend in and seem like they belong in whatever situation they are trying to

perform deviant behavior in; as such, many situations do not involve technical jargon unless necessary.

4.5.7 Impersonal.

[28] discuss how online surveys may seem impersonal. This impersonality may be a strength from the perspective of social engineers maintaining their anonymity. On the other hand, the impersonality of the survey does seem to inversely correlate to the trust of the researcher and the use of the survey itself past the needs of a master's program.

4.5.8 Privacy issues.

Much information has been leaked to the world through the Internet, and social engineers are trained to gather information that does not want to be gathered. However, these social engineers know how to gather information and know that confidentiality is an illusion. Because of this illusion, the social engineers would have to rely on the relative level of trust in the researcher's integrity and the research institution. As such, it was not a surprise when my informant at NIST stated concern about the confidentiality of the survey.

4.5.9 Low Response Rate.

Online surveys "at best attain response rates equal to other modes and sometimes to do worse; and they suggest that the reasons for this merit more study" [28, p. 202]Speculation on this would be that the level of trust between two entities online is less than in person. In fact, beyond using a credit card or debit card to buy items offline in a store, there may be little if any trust transferred between the same two entities that then interact online. This is in some situations contraindicated due to the level of persons giving out information or pictures over the Internet to persons who would not be given this information if met in person.

4.5.10 Word choices.

While researching the lack of data collection of the survey, the researcher saw notes in some sources not to use technical jargon with the general population. This is largely due to the idea that varying backgrounds would make knowledge of specific technical terms not being broadly known. This may have added to the survey not being completed properly as well. The researcher was trying to balance being technical enough to maintain some brevity while not being too concise with the wording of the survey questions. An assumption was made about the widespread knowledge of the jargon used in the survey that should not be repeated; note here that this may make the questions a bit long and therefore present a different reason for not being completed.

5. CONCLUSION

The researcher began by surveying on the Internet. A survey was generated using some Likert-type scale questions, some open-ended questions, a few nominal variable questions, and finally, demographics questions to accomplish this objective. After 2 months of attempting to get answers from the potential 2,246,494 responders, only 12 partial responses were logged. At this point, further research was done as to why the project did not succeed. The researcher recognized the need for interpersonal trust increase to raise the level of data collection. Additionally, research was done on in-group status needs to recruit respondents, and Trust Transfer Theory was discussed concerning the survey project. Finally, there was further discussion about the possible lack of data collection due in part to using an online survey.

The lesson learned here is that the researcher could have received a much higher response rate if there had been a known quantity in the social engineering field. Given the possibility, respondents would have trusted an in-group person well enough to feel confident enough to answer the questions without them being skewed. Lack of trust is part of the nature of social engineers, as their whole existence depends upon taking advantage in some manner of the trusting nature of individuals. The researcher also acknowledges that some of the terms used in the survey are not as widely accepted for their connotation and may not have been understood.

In conclusion, some things could have been done differently with this project. First, the researcher should have taken the time to become part of the social engineering community, as being part of the community would have allowed for a better response to the survey. As a community member, the researcher would have been seen as an individual with a better cost-benefit analysis that would be beneficial to acquiring responses to the survey. This would have been partly due to the persons feeling that they would receive some benefit in the future from assisting with the survey. Finally, the survey should be undertaken again to improve the understanding of social engineering to better grasp the techniques used and improve defenses.

It should be noted that a future researcher could take a few years and become a known quantity in the social engineering world, and maybe even attempt to win the social engineering challenge at Black Hat. Another less attainable option is to survey somewhere on the deep web, but this thesis does not recommend that. A third option would be to get involved with a community of social engineers and do interviews with that group. In this option, the idea would be that doing the interviews would further increase the group's trust and give the interviewer a chance to quell any fears of repercussions directed against the individuals who answered the survey through this method. The interviewer would have to assure the interviewee that they would remain anonymous by not asking personal identifying questions.

REFERENCES

1. C. Hadnagy, Social Engineering: The art of human hacking, Indianapolis: Wiley, 2011.
2. C. Hadnagy, Unmasking the social engineer: the human element of security, Indianapolis: Wiley, 2014.
3. C. Hadnagy, Social Engineering: the science of human hacking, Indianapolis: Wiley, 2018.

4. J. Talamantes, The social engineer's playbook: a practical guide to pretexting, Woodbury: Hexcode, 2014.
5. R. A. Grimes, Hacking the hacker: Learn from the experts who take down hackers, Indianapolis 500: Wiley, 2017.
6. K. Mitnick and W. L. Simon, The art of deception: controling the human element of security, Indianapolis: Wiley, 2003.
7. M. Jakobsson and S. Myers, Phishing aand countermeasures, Indianapolis: Wiley, 2007.
8. F. Salahdine and N. Kaabouch, "Social engineering attacks: A survey," Future Internet, vol. 11, no. 4, p. 89, 2019.
9. N. Galov, "17+ Sinister Social Engineering Statistics for 2021," u.k u.k 2021. [Online]. Available: https://hostingtribunal.com/blog/social-engineering-statistics. [Accessed 19 March 2022].
10. C. Jacomme and S. Kremer, "An extensive formal analysis of multi-factor authentication protocols," in 31st IEEE Computer Security Foundations Symposium, Oxford, 2018.
11. A. Almomani, B. B. Gupta, S. Atawaneh and E. Almomani, "A survey of phishing email filtering techniques," IEEE Comunications Surveys & Tutorials, pp. 2070-2090, 28 March 2013.
12. R. Duncan, "Influence versus manipulation: Understanding the difference," 21 December 2018. [Online]. Available: https://www.forbes.com/sites/rodgerdeanduncan/2018/12/21/influence-vs-manipulation-understand-the-difference/?sh=7947ef9f470c. [Accessed 25 December 2020].
13. E. Maxwell, "Influence vs. Manipulation. Security through education," u.k u.k 2004. [Online]. Available: https://www.social-engineer.org/newsletter/Social-Engineer.Org%20Newsletter%20Vol.%2004%20Iss.%2045.htm. [Accessed 27 Dec 2020].
14. R. Dreeke, "It's not all about "me": the top ten techniques for building quick rapport with anyone," self published, u.k, 2011.
15. M. Karlins and J. Navarro, What ever body is saying: an ex-FBI agent's guide to speed reading people, New York: Harper Collins, 2008.
16. W. E. Saris, Design, evaluation, and analysis of questionaires for survey research, Indianapolis: Wiley, 2014.
17. P. Harris, R. Taylor, R. Thielke, J. Payne, N. Gonzalez and J. Conde, "A metadata-driven methodology and workflow process for providing translational research informatics support," Journal of Biomedical informatics, vol. 42, no. 2, pp. 377-381, 2009.
18. Reddit, "Dive into anything," Reddit, [Online]. Available: https://www.redditinc.com/. [Accessed 22 April 2022].
19. F. K. Wilits, G. L. Theodori and A. E. Luloff, "Another look at Likert scales," Journal of Rural Social Sciences, vol. 31, no. 3, pp. 126-139, 2016.
20. A. Diamantopoulos, M. Sarstedt, C. Fuchs, P. Wilczynski and S. Kaiser, "Guidelines for choosing between Multi-item and Single-item scales for construction measurement: A predictive Validity," Journal of the Academy of Marketing Science, vol. 40, no. 3, pp. 434-449, 2012.
21. E. G. Carmines and R. A. Zeller, Reliability and validity assessment, Thousand Oaks: Sage, 1979.
22. T. R. Hinkin, "A review of scale development prectices in the study of organizations," Journal of management, vol. 21, no. 5, pp. 967-988, 1995.
23. C. E. Neurt and T. Lenzner, "Effects of the number of open-ended probing questions on response quality in cognitive online pretests.," Social Science Computer Review, vol. 39, no. 3, pp. 456-468, 2021.
24. Verizon, "Data Breach Investigations Report," Verizon, 2021.
25. S. Herek, Director, Bill & Ted's Excellent Adventure. [Film]. United States: Paramount
26. R. Borum, The scienc of interpersonal trust, digitalcommons.usf.edu, 2010, p. 574.
27. K. J. Stewart, "Trust transfer on the world wide web," Organization Science, vol. 14, no. 1, pp. 5-17, 2003.
28. J. R. Evans and A. Mathur, "The value of online surveys," Internet Research, vol. 15, no. 2, pp. 195-219, 2005.
29. Cisco, "Vulnerability information. E-mail and spam data," Cisco Talos Intelligence Group, 2021.
30. World Bank, "Individuals using the internet (% of population)," World Bank, 2021. [Online]. Available: https://data.worldbank.org/indicator/IT.NET.USER.ZS. [Accessed 12 Febuary 2021].