Human Exploits in Cybersecurity: A Social Engineering Study

A Social Engineering Field Experiment to Identify Factors for Successful Social Engineering Attacks

By Ian S. Kaufer

A Master's Paper Submitted to the Faculty of the

DEPARTMENT OF MANAGEMENT INFORMATION SYSTEMS

ELLER COLLEGE OF MANAGEMENT

In Partial Fulfillment of the Requirements

For the Degree of

MASTER OF SCIENCE

In the Graduate College

THE UNIVERSITY OF ARIZONA

2016

STATEMENT BY AUTHOR

This thesis has been submitted in partial fulfillment of requirements for an advanced degree at the University of Arizona.

Brief quotations from this thesis are allowable without special permission, provided that an accurate acknowledgement of the source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the head of the major department or the Dean of the Graduate College when in his or her judgment the proposed use of the material is in the interests of scholarship. In all other instances, however, permission must be obtained from the author.

SIGNED: Ian Kaufer

APPROVAL BY MASTERS PAPER ADVISOR

This thesis has been approved on the date shown below:

_____ 04/11/2016

Dr. Advisor Name Date

Advisor Title of Management Information Systems

ACKNOWLEDGEMENTS

I would like to thank my research partner, Brendan McDermott for his guidance, wisdom, and friendship throughout this research project, independent studies, and master's process. I would also like to express my gratitude to my advisors, Dr. Jesse Bockstedt and Dr. Matthew Hashim for their inspiration and persistence in making this research project a reality. I would also like to thank Cathy Larson, Dr. Mark Patton, and Dr. Hsinchun Chen for their incredible support and guidance in this Scholarship-for-Service program throughout the past two years. I would also like to thank all of the following students who came out in the field with us in the data collection phase: Cyrus Afarin, Calvin Barreras, Vincent Ercolani, Tiffany Feller, Ashley Ireson, AJ Jicha, Dominic Kaufer, and Jasper Puracan. Finally, I would like to give a special thank you to my parents, brothers, sister, dog, and friends for their continuing support of my personal and professional life goals.

Table of Contents

List of Figures

List of Tables

Abstract

Purpose

Social engineering is an information security threat that continues to plague organizations today. As much as organizations can invest in technical security products and services to protect their networks, the human element in security is weak. Social engineers are malicious attackers who exploit the vulnerabilities in human behavior to gain access or retrieve information. In the realm of information security research, there is quite a lot of research on technical security products and services. However, there have been no direct, field experiments to test the factors that make social engineering more or less successful in a physical, non-technical environment. The purpose of this report is to discuss the details going into our social engineering experiment, the Institutional Review Board (IRB) process, literature review, experiment and design, hypotheses, analysis, motivation, and discussion for why we conducted this research.

Design

Using a unique field experiment, subjects participated in research by being approached proactively by confederates who manipulate factors such as social context and reward. The data was gathered by confederates who used a survey questionnaire to collect a subject's information relating to personally identifiable information (PII). Researchers on-site recorded whether or not subjects responded to certain questions, while also conducting debriefs with the subjects in order to request consent to use the responses of the surveys for analysis.

Findings

The analysis of results shows statistically significant responses by subjects for a certain combination of social context and reward. For example, subjects were more likely to divulge more sensitive information from the survey questionnaire when the subject was presented with a confederate from a charitable organization paired with a higher reward. Other findings showed that subjects were more willing to provide more severe PII responses when paired with the divulgence of the last five of the social security numbers. Based on the quantitative factors that were analyzed, there are significant findings that could have further implications towards later research, and practical applications.

Originality

The experiment design has not previously been implemented with the design for social engineering field experiments, although with IRB approval, this design can be reproduced to test other factors relating to social engineering. This experiment allows researchers to test social engineering in a real-world environment where subjects are not aware of any research being conducted. This allows the researchers to not bias their own results.

Introduction

Social engineering (SE) is a major issue that organizations are having trouble dealing with today. By using behavioral vulnerabilities instead of technical ones, hackers can rely on tricking humans into providing information when the victim doesn't expect it.

The idea of social engineering is present in many areas of information security, as well as cybersecurity. In the context of computer security exploitation, the reconnaissance phase of security-testing and penetration-testing exercises can be conducted using social engineering in order to obtain initial information that can be leveraged later on in the exploit process. While computer security exploitation can utilize the social engineer's expertise, information can be exploited through the use of non-technical, behavioral security exploitation.

There are a many attack surfaces in which social engineering can occur. In many cases, social engineering is used to commit identity theft/identity fraud, by which an attacker assumes the identity of a victim after obtaining personally identifiable information about them (Peltier 2006). These situations can have major implications for the victims; over 16.6 million Americans were identity theft victims in 2012 – the effect of identity theft can be felt for years and are "more likely to experience financial, credit, and relationship problems and severe emotional distress" as a result of social engineering (Justice 2013). Knowing the difficulty in which these attacks can have an effect on citizens can make one realize the severity in which social engineering attacks have on everyday life. Other social engineering attacks can be found during the information gathering/reconnaissance process for a business, personal attacks, and other malicious events.

For businesses, the threat of social engineering attacks are high. Of the larger businesses surveyed in a study, 48% fell victim to social engineering attacks (Dimensional Research 2011). Not only are the cost in losses significant for each social engineering incident, but they are also reoccurring, costing the business more in responding and countering incidents in later times.

The anecdotal resources for these attacks and the nature in which social engineers conduct their attacks is mostly from evidence in fictional literature, social events, and non-academic experiments: Kevin Mitnick's book The Art of Deception, DEFCON's Social Engineering Village (event), and The Robin Sage Experiment (Goodchild 2010).

The problem is that there have not been direct academic investigations into what conditions, factors, and situations make social engineering in a physical environment work and why. The Software Engineering Institute has conducted a review of what human factors may cause social engineering attacks to be successful, but have not applied those factors to a field exercise (SEI 2013). Understanding if, and what, factors can make individuals more or less likely to provide their personally identifiable information (PII) is important for future security training and awareness, and continuing research in this field. By conducting a field experiment in the social engineering field, this research can generate discrete data for conducting formal statistical analyses and further aid in the design of social engineering experiments in the future. It can also provide businesses, business leaders, and security experts with information of what makes social engineering possible, as well as potential countermeasures in the formation of policy, training, or defenses. Through the course of 3 semesters between spring 2015 and fall 2016, Brendan

McDermott and I began to delve into social engineering literature, experiments, designs, methodologies, and eventually came up with a research question that could be tested for social engineering dilemmas.

People may be more likely to be socially engineered when factors such as context, reward, and gender are combined or altered. The goal of this Master's report is to identify if these factors make social engineering more or less likely to be successful, as well as how those factors affect the amount of personal information provided by the victims in these attacks. It also discusses the process for conducting Human Subject Research in this context – where subjects are unaware of the research being conducted in actuality. It begins with an overview of social engineering and the context in which our research focuses on, then moves into a literature review of related research, followed by a section for our model/hypothesis. Following the hypothesis is a section for our motivation to conduct research in this field, the methodology and design of the field experiment, an analysis and results section of our findings, then a conclusion of our research. This analysis is broken down into qualitative and quantitative analyses, since there were significant elements of the study that help further the research goals and findings. While both quantitative and qualitative aspects were captured during the experiment, the quantitative aspects are detailed in this paper. Brendan McDermott observed and detailed the analysis further for qualitative results during and after the experiment.

Literature Review

Social engineering is based in a few categories that each provide relation to how and why people are subject to social engineering-based attacks. These categories include social psychology, behavioral economics, and information security (Ross 2008). The prior literature for social psychology focuses on how emotions and cognition effect peoples' decision-making ability, given certain factors. Social engineering has a basis in behavioral economics through research on how people make decisions in social contexts, while also identifying game theory and economic-related decision making aspects. Marketing studies help to identify how people are attracted to items through direct marketing, which provides value for a physical social engineering experiment. Finally, information security has recently began focusing more on how the human element of security can be the weakest link and easier layer of security to exploit. Synthesizing each area of literature in the context of social engineering research provides a lens for examining what's been researched before, what hasn't been done, and the research gap we can fill to the best of our knowledge. The literature review will be presented by the fields of research, then finish with a synthesis that leads to the research gap.

Social Psychology Literature

Social psychology research helps build a foundation for social engineering and deception because of the fact that people become victims of social engineering attacks through the decisions they make. Decision-making and cognition research is an area of research that social psychology has focused on.

Decisions are made after options are considered, although the process taken to make a decision is made through a quick reasoning strategy (Gilovich, Griffin, & Kahneman, 2002). This quick reasoning strategy may be exploited by a malicious social engineer by identifying the option that is most luring. If there are external factors involved, such as the lure of a reward after completing a survey that includes the release of personal information, this can lead people towards a certain direction (Plous 1993). Furthermore, the introduction of risk in a situation will affect how a decision is made – either by a quick reasoning strategy, referred to as "the affect heuristic" or by doing analysis on the situation (Slovic, Peters, MacGregor, Finucane 2005). Other aspects of human psychology lend to how people make decisions, such as the "interplay between emotion, cognition, and decision making" (Schwarz 2000). Schwarz notes that emotion and mood related to the context of the decision that needs to be made. For example, if a person is presented with a survey from an organization that represents a cause the person believes in, the "congruent" nature of the cause with the person's beliefs lead to a decision made by emotion and mood.

Behavioral Economics Literature

Not only do social engineers look towards how human nature and decision-making, but they also use items of monetary value as an incentive to lure unsuspecting victims (Applegate, 2009). Grossklags et al. (2007) looks at a subject's willingness to pay for protecting their personal information, as well as a subject's willingness to accept payment for their personal information. According to Grossklags, even though some people are aware of the sensitivity of their PII, they are still willing to sell / trade off some of it or give it away based on convenience or bargain for a particular reward. In another study, Grossklags and others came to the conclusion that people will trade off their PII for short-term benefits / rewards (Grossklags 2005). They looked at three

main themes for decision-making by subjects, and the constant throughout the themes pointed towards a consistence of subjects to "deviate from the rational strategy", meaning that subjects incorrectly predict how decisions made now may affect their future. Even with full information, people may still make decisions that don't provide the optimal benefit.

Other studies have been conducted to specifically identify the level of risk aversion as the payment amount scales up or down from making choices in a lottery. As one main focus of our research includes using real incentives as part of the social engineering design, prior literature in this area provides exceptional value in completing the quantitative analysis based on this factor. By treating the high and low lottery reward amounts in isolation, the results of their research show that subjects become increasingly risk averse as the payoff scale increases (Holt and Laury 2005). The findings section may show (need to see if this is the case…) a congruency between Holt and Laury's research and our data.

Another main focus of our research includes using social context as a treatment variable in order to test the two factors in the context of a social engineering study. Context as a factor plays a large role as context changes (Colin and George 2004).

As a note, our research uses incentives based on reward in high and low categories, although the incentive itself is not cash. The incentives our experiment uses are items valued high and low. Comparing monetary versus item incentives was out of scope for this experiment.

Information Security & Social Engineering Literature

Social engineers prey on the qualities of human nature, such as the confirmation bias in assuming strangers can be trustworthy, wanting to be helpful towards a good cause, and the desire to receive something quicker without having to pay for it (Peltier, 2006). There are three key aspects that are highlighted in Peltier's paper regarding how social engineers can deceive the subject: using alternative routes to persuasion, exploiting attitudes and beliefs that affect the interaction, and leveraging techniques for persuasion and influence (Peltier, 2006). It is these qualities and persuasive tactics that social engineers may favor as items to exploit in order to obtain information.

Interestingly, there have not been any direct experiments to test the factors which make social engineering possible. To the best of our knowledge, only a few researchers take a direct, field experiment approach to social engineering. One study socially engineers employees of an organization through a questionnaire provided in an email link (Workman 2007). There has also been a social engineering experiment to test the antecedence to being susceptible to social engineering by using social context, which was also conducted through email (Jagatic 2007). Both Workman and Jagatic et. al. conducted these experiments after first working with their Institutional Review Boards to obtain Human-Subject Research approval, similar to our experiment design.

Other prior research using a more direct field experiment includes a study conducted through a non-personal, non-physical means, via a phone call (Müller 2009). As much research has been done in these kinds of direct field experiment settings, our focus was to fill the gap in social

engineering research by applying a direct, in-person, social engineering exercise against unwitting participants.

Research Gap in Literature

While social engineering literature is diverse through the fields of research identified, there is a lack of research that combines these fields in an inclusive study in the social engineering realm to not only test the physical elements of social engineering exercises, but to also further understand what factors make people more likely to give up their personal information. The difficulty in creating a field experiment in a physical environment with deceptive design creates unique challenges.

When the physical, in-person element of social engineering experiments coming into play, the difficulty in designing an experiment that conforms to the rigor and acceptance of the Institutional Review Board increases. Jakobsson identified the main difficulties for obtaining IRB approval in human-subject social engineering experiments, by looking at the difficulty in dealing with ethics and regulation, psychological harm to subjects following experiment debrief, and legal issues regarding personal privacy information (Finn and Jakobsson 2007). As our research moved forward, we also came across a few of these difficulties, as the IRB at the University of Arizona found it difficult to accept a proposal that purposely deceives subjects without prior knowledge of a study being conducted. This is another gap that our research aims to fill by designing an experiment that the IRB can accept, as well as designing a methodology that allows for both quantitative and qualitative statistical analysis.

The two main focus areas that our research fills are: identifying what factors make social engineering more likely to be successful, and an analysis of conducting a social engineering field experiment in-person. Through the unprecedented acceptance of the design and methodology of our social engineering experiment by the University of Arizona's Institutional Review Board, our research has been granted access to this yet studied area of social engineering research. Having solved these challenges through proper channels of research design, development, and human research ethics covered, we were able to develop a study to test the prior literature, as well as test our hypotheses regarding social engineering susceptibility.

Motivation

This paper and our research aims to see if people are altruistic when presented with differing factors, such as context and rewards. For example are people more willing to give information because they think the more information they divulge, the more it will help Books for Kids foundation or BNI Market Research firm. In the Books for Kids scenario, does the prospect of filling a survey for a charitable organization result in the subject feeling joy or altruistic feelings.

Our interest also lies in the behavioral aspects of decision-making when presented with incentives to participate in responding to a questionnaire, paired with the hypothesis related to altruism. Understanding how context relates to altruism, while paired with incentives and the subject's willingness-to-sell is a motivation behind this research (Grossklags, 2007).

Specifically looking at the social engineering attack was also a motivation during the initial stages of this research. Looking at what PII can be severely damaging to the individual, as well as what PII is utilized in normal day-to-day can be a gold mine for social engineers. By reviewing password reset forms and its questions, credit card application questions, and general PII from surveys, our questionnaire categorized questions based on the damaging nature of specific questions, as well as the importance from clusters. Social engineers generate profiles of targets, and thus having the profile information from subjects in the experiment helps to build a credible social engineering platform for study.

Aside from the behavioral and psychological aspects that motivate this research, there was also significant motivation in understanding how to design and implement a field experiment with deceptive constructs. Understanding the proper IRB protocol and ensuring human subject participation and safety has been detailed for other phishing experiments previously (Finn 2007).

Purpose

The purpose of this research was to investigate what factors make social engineering more likely through context and reward constructs, as well as understanding how to design an experiment that deceives subjects with IRB approval. The addition of these two themes in our research aims to contribute new knowledge to precursors of human deception, as well as provide an analysis of captured data for study and future research. The real-world implications include policy-level guidance for how to better train and protect sensitive information. In the context of practitioners

dealing with real social engineering threats, understanding the basis for those threats can help an organization's security posture through more informed training and awareness.

Hypotheses

Our hypotheses were motivated by our assumptions for behavior when presented with information and the potential for economic gains. After the literature review and a discussion about how we could implement the design of the experiment, we wanted to ensure we were testing appropriate factors. Using the results from Grossklags on people's willingness to sell their information, we concluded that people will give up more information from our study if they were presented with a combination of factors that allowed for a greater earning amount (Grossklags 2007). That is, people earn more when receiving a high reward versus a low reward. People also "earn" with respect to altruistic gain when they feel like they are doing something of social benefit, or the greater good. By aligning our main effect factors with elements from Grossklags' paper, we identified two main effect hypotheses, as well as an interaction effect hypothesis that may provide insight into successful or unsuccessful factors to exploit.

The following hypotheses are the general research questions we were interested in answering, with more detail and analyses of results in the following section.

Hypotheses 1 – Main Effect

People are more willing to provide more PII when filling out a survey for a charitable organization, as compared to a market research firm.

Hypotheses 2 – Main Effect

People are more willing to provide more PII when filling out a survey for a high reward, as compared to a low reward.

Hypotheses 3 – Interaction Effect

People are more willing to provide more personally identifiable information (PII) when filling out a survey for a charitable organization and for a high reward, as compared to a market research firm and for a low reward.

Experiment Design and Methodology

Our research experiment was designed to capture information from subjects in the field. Using the two factors of context and reward, we used the following experiment design and methodology to test our hypotheses.

In order to move forward, we had to determine logistics for where we could complete the data collection, what times and days of the week, as well as operational details such as confederates to work the field with subjects. Confederates would approach subjects while wearing context badges and offer raffle rewards following the subject's participation to fill out a survey. Before setting out into the field to begin the experiment, other hurdles had to be met, including understanding what the survey questions will be and why, as well as obtaining approval from the university facilities management to conduct a survey, and obtaining IRB approval to carry out the experiment.

Design

During the initial phase to identify prior literature and research in this area, Brendan McDermott and I began to consider what factors we could manipulate in a field experiment of this type. Considering the subject's initial greeting with a confederate, we came to the conclusion that the first observable interaction between a confederate and a subject would include the confederate explaining who he/she was with reference to the organization they are a part of. This became the context factor (market research firm vs. charity organization). Looking towards the motivation for how individuals make decisions when presented with payoffs, we also identified elements of

behavioral economics using raffles for payoffs. This became the reward factor. High versus low rewards have differing effects on decision-making, as discussed in the prior literature section for behavioral economics. In order to have statistical significance when evaluating results, the 4 possible combinations of the 2 factors would have to produce a total of 120 data points. The breakdown of data points suggested is listed below in a table.

	High Reward	Low Reward
Profit Organization	30	30
Nonprofit Organization	30	30

Table 1: Recommended Data Points and Breakdown of Factor Combinations

The breakdown includes the combinations high reward and profit organization, high reward and nonprofit organization, low reward and profit organization, and low reward and nonprofit organization.

In order to test social engineering tactics and obtain true and unbiased results from subjects, the design includes a fair amount of deception against the subjects. The context and survey purpose is initially described to subjects while offering a potential reward through a raffle. Once the subjects completes the survey, the confederate walks the subject to the researcher completing debriefs, where the true nature of the survey is revealed to the subject. Given the deceptive nature of this experiment, a formal process and review must be completed to ensure the subjects of the experiment are not being unduly harmed mentally, nor psychologically. This process

includes a formal evaluation of research design as well as requirements for how to comply with human subject research policies and guidelines. This was completed by the Institutional Review Board, or IRB.

The process from start to finish in our experiment includes the data collection phase where confederates obtain information from subjects, the transformation phase where researchers ensure anonymity, then destruction of personal data, and finally the analysis of results phase. The three phases in this process can be seen in the figure below.

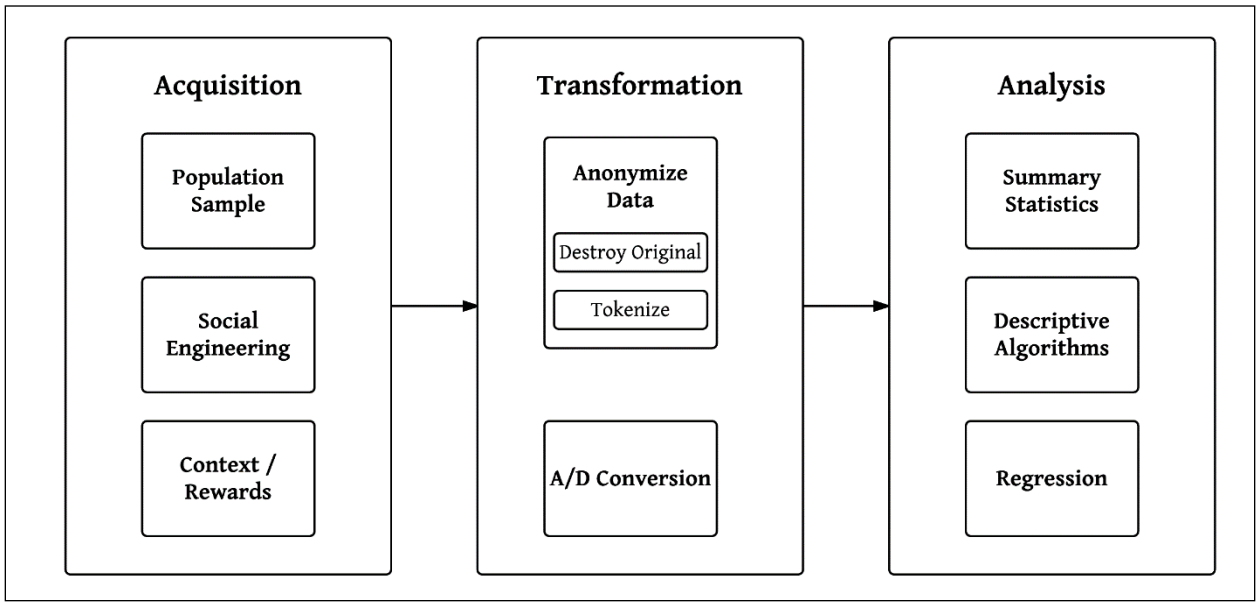


Figure 1: SE Experiment Phases

In the acquisition phase, confederates would identify subjects (population sample) and use social engineering based on context and reward factors. Once the subject has completed the survey and

finishes the debrief with a researcher, another researcher on-site transforms the analog data by digitally recording responses to questions, as well as destroys the original data provided by the subject. Tokenization in this process refers to the sanitation process of anonymizing any recorded information. Finally, the analysis phase includes the process by which researchers analyze whether the results were significant or not.

Methodology

The methodology of this experiment can be broken down into multiple parts. This includes the operational aspect of obtaining approval to conduct the study, the experiment and how the design would be implemented, and the process of completing the social engineering experiment itself.

Operational Aspects

Operational aspects include information for: obtaining IRB approval, researchers, confederates, factor combinations, field location, time of day and week, environment conditions, researcher-confederate layout, average confederate-subject session time, and the process for conducting the experiment.

Obtaining IRB Approval

In order to conduct this experiment, our design required human-subject approval by a review board before beginning data collections. The process to obtain this approval lasted approximately 3 months from submission of forms to being granted access to conduct human-subject research. The process included submitted multiple lengthy forms including: F200-Application for Human Research, F107-Verification of Human-Subject Training, Estimated Costs (to conduct

experiment), and Backstory for Confederates relating to social engineering Contexts, Debrief Script, and Frequently Asked Questions (FAQ) for Confederates (to respond to questions from subjects). The process also included approval from Facilities Management to utilize physical area to conduct surveying.

In order to ensure human-subject safety for the entire project, researchers and confederates were also required to train on the history, ethics, and guidelines for conducting any research involving human-subjects. This training included Collaborative Institutional Training Initiative (CITI) program material for Social and Behavioral Research Investigators, as well as Arizona-specific training in Native American Research.

Researchers

Generally, the experiment included 4 researchers; two faculty advisors and two associate researchers. The two faculty advisors were Jesse Bockstedt and Matthew Hashim. The two associate researchers were Ian Kaufer and Brendan McDermott. Brendan and I guided confederates in the experiment, as well as performed operational details such as deciding start and end times for each round, data transformation, subject debriefs, and all other details during data collection.

Confederates

During the data collection stage, there were a total of 8 confederates contributing to the study. Confederates interacted with associate researchers and subjects during the experiment. Confederates were mainly Master's in Management Information Systems (MIS) students from

the Scholarship-for-Service (SFS) program, although there were three undergraduate confederates from varying majors and academic terms (based on freshman, sophomore, junior, and senior). The confederates were recruited by the researchers and provided compensation for their time and efforts in the study.

Factor Combinations

In order to ensure the results could be analyzed with enough statistical significance, our research advisors recommended obtaining at least 30 recorded observations for each factor combination. Our factors included context (market research firm vs. charitable organization) and reward (high vs. low). With four factor combinations possible, a total of 120 recorded observations was required. With 10 total sessions, our data collection concluded with 118 recorded observations, which was determined as sufficient by the faculty researchers given the closeness to 30 observations per group.

Field Location

The area in which the field experiment took place was a high pedestrian-traffic area at the University of Arizona. The specific location was a seating area just north of the Student Union Memorial Center, between a parking garage and Student Union shops. Given the proximity to campus buildings for lecture, work, and the Student Union, the strategy included utilizing an area where many subjects are likely to be traversing. The higher the number of potential subjects crossing an area increased the chances that confederates could approach and begin the social engineering experiment, which is pictured next. The location used was next to the Student Union at the seating area facing north.

Figure 2: Seating Area Used for Field Location

Time of Day and Week

The time of day chosen for the majority of sessions was during the afternoon, between 12:45pm and 5:30pm. This time was chosen based on the strategy that subjects will be likely to be finishing lunch, leaving class, or walking the campus area during that time. The day of the week for when the sessions occurred were on Tuesdays, Wednesdays, and Fridays. The day of week chosen was mainly based on open schedules for researchers and confederates involved. There were a total of 10 sessions in which researchers and confederates went into the field to complete the data collection portion of the experiment.

Environment Conditions

The environment conditions were between 65-80 degrees Fahrenheit with blue skies on most days, or overcast skies on other days. One data collection day was cut short halfway through when it began to rain.

Researcher-Confederate Layout

The logistical layout included one researcher in a corner responsible with recording and shredding PII, while another researcher responsible with debriefings was approximately 20 feet away at a table with 3-4 chairs. The debriefing researcher had visibility to both the other researcher, confederates, and subjects. The confederates stood approximately 15-20 feet away from the debriefing researcher, while in closer proximity to potential subjects walking by.

Average Confederate-Subject Session Time

Based on the first session with a 2-researcher and 1-confederate layout in the field, each 2-hour session would conclude with approximately 12 recorded observations from subjects. Approximately 20 hours was spent in the field conducting data collection by the two associate researchers.

Process

The process for conducting the experiment itself includes meeting with confederates, conducting survey collections, debrief participants, recording and destroying data, and distributing raffle rewards.

Meeting with Confederates

Two researchers would meet with the confederate(s) and carry the materials and paper shredder to the field location approximately fifteen minutes prior to commencing the experiment. The other researcher and I would prepare the confederates for the four sessions they would complete in the field by giving them badges and choosing a reward in which they would present to the subjects. An example badge that a confederate would wear is shown below, where the subject's name is listed, as well as the context for the session.

Figure 3: Badges for Confederates with Context Noted

Confederates would walk to their layout position and begin the experiment to socially engineer subjects as they walk by. Each factor combination represents a session, and so four sessions per hour and a half equates to one session every 20-25 minutes. In order to control the confederate and their approach to subjects as much as possible, we compensated the confederates for their time as well as reminded them that the surveys they collect is not a competition against other confederates.

Conducting Survey Collections

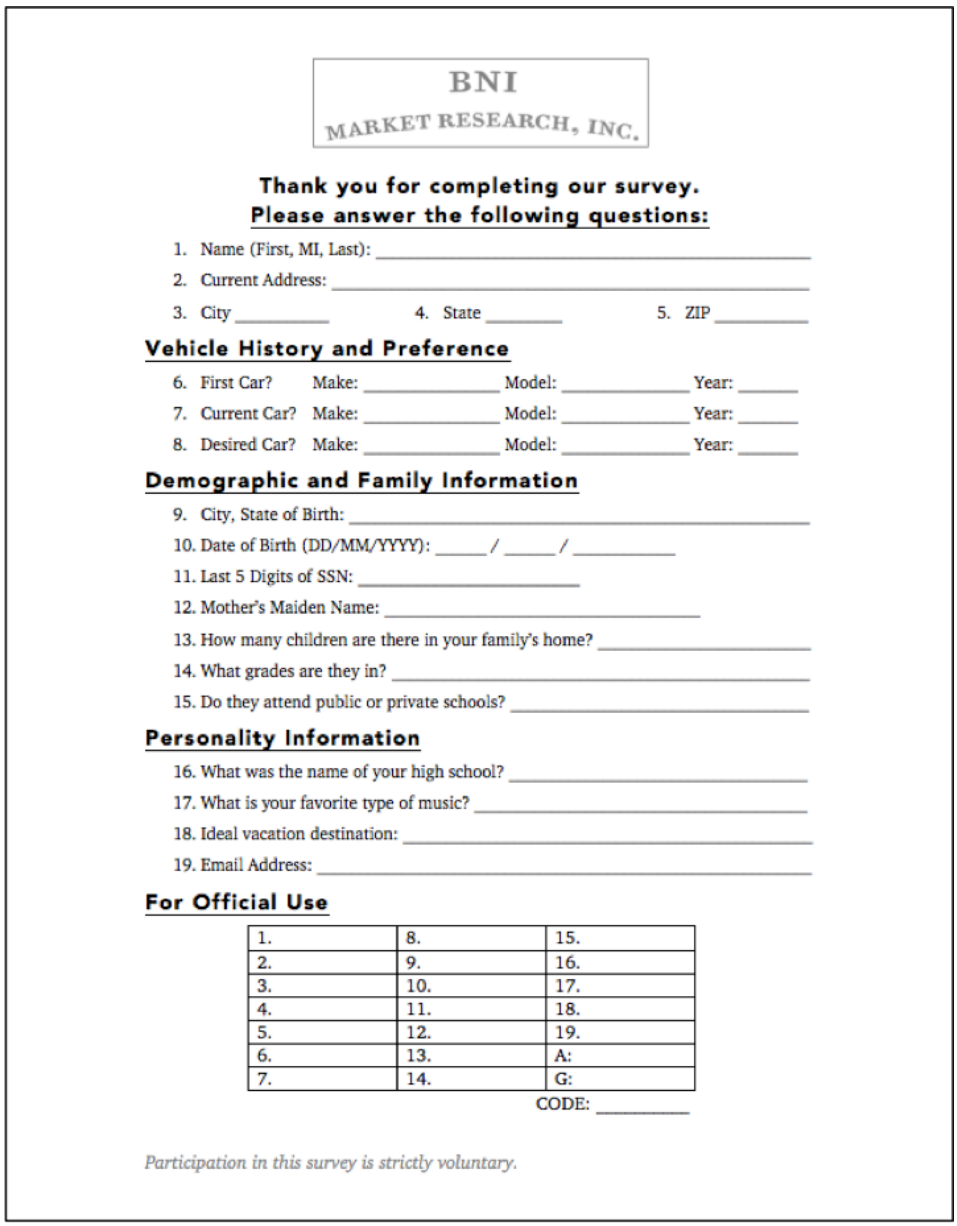
In order to control for a non-deterministic selection of subjects, the confederate would proactively approach every 3rd potential subject who passes by and begin with the script for introducing themselves, the context, and reward. The script used by the confederate for approaching subjects is located in Appendix A.

Subjects that responded to confederates would use the pen and clipboard to fill out the survey in which the confederate was holding. In order to maintain IRB approval, if the subject questioned

the survey then the confederate would reassure the subject that all questions are optional, and the raffle can be entered without having to respond to all questions. The confederates would review the Frequently Asked Questions document provided during the confederate preparation, in the case that subjects wanted more information regarding the context specifics or reward details. This FAQ document is located in Appendix B.

Survey questions ranged from demographic information, personal history, and sensitive personal information, although the sensitive information was not transparent to the subject. There were 19

questions on the survey, and an example can be seen next. This example shows the BNI Market research firm logo at the top, signifying the commercial, for-profit organization as the context.



BNI
MARKET RESEARCH, INC.

Thank you for completing our survey.
Please answer the following questions:

1. Name (First, MI, Last): _____
2. Current Address: _____
3. City _____ 4. State _____ 5. ZIP _____

Vehicle History and Preference

6. First Car? Make: _____ Model: _____ Year: _____
7. Current Car? Make: _____ Model: _____ Year: _____
8. Desired Car? Make: _____ Model: _____ Year: _____

Demographic and Family Information

9. City, State of Birth: _____
10. Date of Birth (DD/MM/YYYY): _____ / _____ / _____
11. Last 5 Digits of SSN: _____
12. Mother's Maiden Name: _____
13. How many children are there in your family's home? _____
14. What grades are they in? _____
15. Do they attend public or private schools? _____

Personality Information

16. What was the name of your high school? _____
17. What is your favorite type of music? _____
18. Ideal vacation destination: _____
19. Email Address: _____

For Official Use

1.	8.	15.
2.	9.	16.
3.	10.	17.
4.	11.	18.
5.	12.	19.
6.	13.	A:
7.	14.	G:

CODE: _____

Participation in this survey is strictly voluntary.

Figure 4: BNI Social Engineering Survey

The turnover time for each confederate-subject interaction remained relatively constant at 3-4 minutes. This included the amount of time to fill out the survey, ask questions of the confederate, and be walked over to the area where debriefs would occur with a researcher. In order for the confederate to ensure the subject would walk over to the researcher for the debriefing, the confederate would say the following:

"Please see my supervisor for a raffle ticket. He is seated right over there with the documents on the table. It will only take a minute or so."

Debrief Subjects

The researcher conducting debriefs would ask the subject to be seated, and that it will only take a few moments. The researcher would begin debriefs by describing the experiment, as detailed in Appendix C. This script was used to ensure debriefing subjects would remain constant. An example table with a researcher seated is shown below.

Figure 5: Debrief Table with Researcher Seated

The researcher would provide the following information during the debrief: thank the subject for taking the survey and participating in the experiment, provide information regarding the experiment and why it was conducted, inform how the survey itself would be destroyed using a paper shredder, faculty advisor contact information, and ultimately request consent from the subject to maintain the perforated section below the survey for research and analysis. The average time to complete one debrief varied, since there were times when the researcher had to complete multiple debriefs at the same time with subjects arriving from multiple confederates in the field. Generally though, each individual debrief lasted approximately 4 minutes.

The course of debriefing included obtaining consent from the subject, recording the subject's information on a raffle entry form if the subject chose to be in the raffle, and allowing the subject to field any questions during the experiment. Since the original survey with subject information was to be shredded, the raffle entry form was provided as an actual document to be kept for the raffle drawing later on. This form included two form fields: name and email address.

Recording and Destroying Data

The bottom section of the survey was "For Official Use", with a perforation above it. This was the section that would be torn from the top half, in order to record whether or not the subject

responded to a question or not. If the subject consented, the perforated section would be kept; if the subject did not consent, the perforated section would be shredded.

Once the subject consented to providing the perforated section of the survey to us for research, Brendan would store the perforated sections of the surveys in an envelope to be recorded later. An image below shows Brendan waiting for confederates to deliver the subject's survey for recording and destruction.

Figure 6: Researcher Waiting to Record/Destroy Survey Responses

Following the survey sessions, Brendan and I would transform the analog, perforated sheets into digital records through spreadsheets on a computer. This would allow for easier manipulation for charts, graphs, and eventual regression analysis.

The section of the survey that contained the subject's information was shredded on-site with a paper shredder following every survey that was filled out. The shredded strips of paper were stored in a bag until the data collection period was over. The final step in confirming shredded PII was in fact destroyed included taking the shredded material to a secure trash bin in the MIS Department.

Conducting and Distributing Raffle Rewards

In order to ensure IRB approval for the deceptive experiment, the raffle element had to be real. Following the study, Brendan and I selected 2 raffle winners: one for the high reward (iPad mini) and one for the low reward (pizza coupon). Using Excel and the RAND function, two subjects were selected. With final funds approved, the raffle items are being purchased, and the subjects are being contacted for selection. If the subjects do not respond within a 10 day period to collect their rewards, we would select two more subjects and random from the list of those that entered the raffle.

Analysis and Results

The following analyses are divided into quantitative analysis and qualitative analysis. The quantitative analysis is the main focus of this report, whereas the qualitative analysis is the focus of my research partner's report, *Factors Enabling Fraud: A Study of Social Engineering and Identity Theft* (McDermott 2016).

Quantitative Analysis and Results

The quantitative analysis is broken down into two parts: summary statistics and regression analysis. The summary statistics section describes counts, percentages, and demographic information about subjects and what they responded to. In the regression analysis section, we present our model for completing a logistical regression with multiple dependent variables and the assigned covariates for measuring statistical significance.

Summary Statistics

Total Responses, Rejections, and Interactions

The outcome of our data collection effort resulted in 118 total surveys collected. With 540 total observed rejections from the 8 confederates, there were a total of 658 interactions between the confederates and subjects. The confederates had, on average, a 17.9% success rate in having subjects respond to surveys and ultimately consent to participate in research. The chart below shows the total number of responses and rejections, along with percentages of the total interactions.

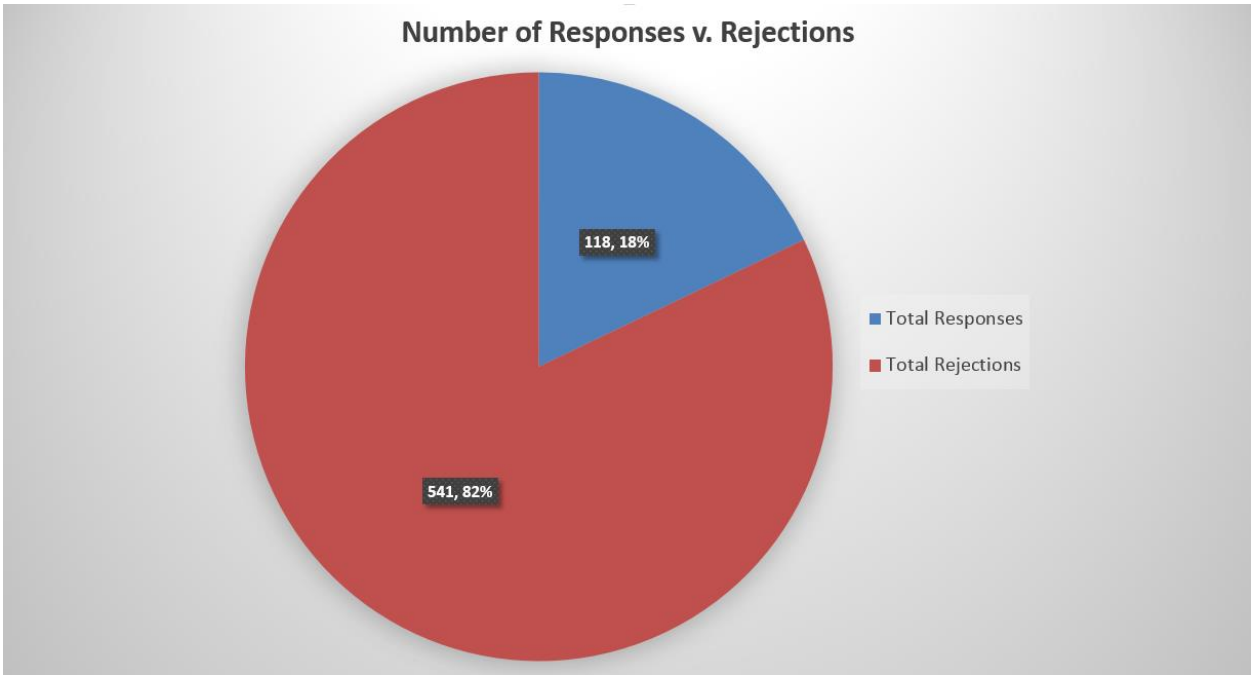


Figure 7: Survey Responses v. Rejections

Response Rate by Confederate

Interesting to note was the difference in successful responses by confederates. In total there were 8 confederates that assisted in our experiment. While there were controls during the experiment (badges, general scripts, FAQ, context, reward), there are a number of other potential confounds that could contribute to the differences in successfully luring potential subjects to take the survey. The following table shows the difference in success versus rejection by each confederate.

Confederate	Rejection Count	Successful Survey Count	Total Count	Success Rate
Calvin	33	88	121	27.3%
Dominic	11	50	61	18.0%
Ashley	39	180	219	17.8%
AJ	13	61	74	17.6%
Cyrus	6	30	36	16.7%
Tiffany	5	35	40	12.5%
Jasper	7	53	60	11.7%
Vincent	4	43	47	8.5%
Average	**14.75**	**67.5**	**82.25**	**17.9%**

Table 2: Success Rate by Confederate

Looking at the average responses by all confederates, they performed with 18% success. However, there is a relatively large range between the confederates. Only two of the confederates, Ashley and Calvin, participated one more day each than the other confederates. The greater success of those two could be attributed to confound of greater success over time due to practice as to what works and what does not. Other confounds for the confederates could also include personality traits, confidence, and experience in conducting surveys. My research partner, Brendan McDermott utilizes a qualitative research framework to identify themes and traits across confederates for a more full understanding of the greater context surrounding this experiment.

Total Responses by Factor Combination

With 118 surveys collected determining statistical significance requires approximately 30 observations for each factor combination. There were four factor combinations, requiring a total of 120 observations. The table below shows the number of recorded observations by context/reward combinations, as well as the percentage of responses for the combination out of

the total observations. According to Dr. Bockstedt, the breakdown of factor combination responses was enough for statistical significance.

	iPad Mini (High Reward)	Pizza Gift Card (Low Reward)
Books For Kids Foundation (Charity Context)	29 (24.6%)	33 (28%)
BNI Market Research Firm (Market Research Context)	27 (22.9%)	28 (23.7%)

Table 3: Four-factor Combination Responses

Total Responses by Survey Question

The distribution of answers also shows the overall significance of what subjects are willing to provide in a survey. The following chart shows total responses based on question, irrespective of the factor combination, for demonstration that people will provide information.

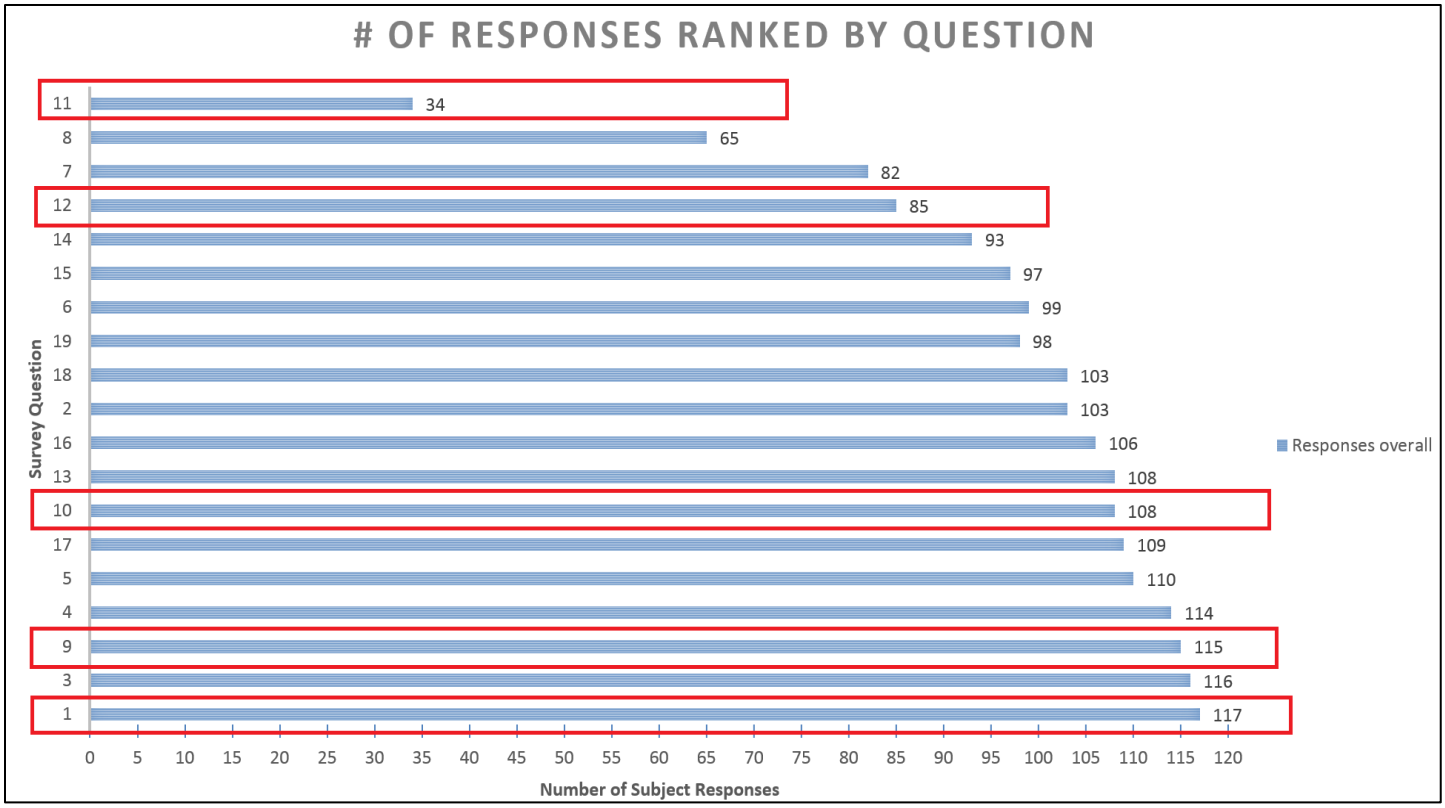


Figure 8: Count of Survey Responses by Question

Most notable in the chart are the number of responses to questions we defined as personally identifiable information (PII). PII in this study is defined by the following five items located on the survey: Name (Q1), City and State of Birth (Q9), Date of Birth (Q10), Last 5 of the Social Security Number (SSN) (Q11), and Mother's Maiden Name (Q12). By only observing the responses overall, subjects were willing to divulge an astounding 28.8% of their PII, since everyone who divulged their SSN (Q11) also provided all of the other four PII questions on the survey.

Responses to PII by Factor

One of the main drivers for this research was to observe the response effect of requesting subjects to fill out a survey containing (PII). Assumptions going into the study included thinking that subjects would be aware of their PII and not divulge it under any circumstance. That assumption was proved to be invalid after analyzing results from the study. The following table shows the percentage of responses to PII when presented with the factor combinations, as well as the total percentage of responses to PII when looking at only one reward element or one context element.

		Context		
Reward		Charity	Commercial	Total
	Pizza	31.0%	40.7%	35.7%
	iPad mini	22.2%	15.4%	19.4%
	Total	28.3%	26.2%	27.1%

Table 4: Percentage of PII by Factor Combination

Interestingly, looking at the low reward alone, subjects were more willing to provide PII than the high reward alone – Nearly 36% of subjects provided PII when the potential for reward was pizza, while only 19.4% of subjects provided PII when the potential for reward was an iPad mini. As for the context factors alone, subjects were slightly more likely to provide PII when presented with a charitable organization such as the Books-for-Kids Foundation, as opposed to the market research firm, BNI Market Research. In that case, the difference was less obvious, although 28.3% of subjects provided PII for the charity context, versus 26.2% of subjects who provided

PII for the commercial context. Overall, based on both factor combinations, 27.1% of subjects provided their PII.

There are several reasons we discussed during the analysis phase as to why more subjects would provide PII when presented with the low reward. The closeness of the field location to the Student Union could present a situation akin to the "time inconsistency" problem from behavioral economics where the immediate gratification of a potential pizza reward is directly correlated to the pizza provided in the Student Union eating area. Another reason could have been the responses by subjects that they don't need another iPad since many noted already having one. While this is contradictory to the notion of economic gain, since an iPad is worth significantly more than a pizza in terms of dollar-value, we as researchers assumed the subjects would understand this distinction. Along with this notion is also the fact that the field location for this experiment was on a college campus with the majority of subjects between the ages of college students. While there is the possibility that not all subjects were college students, many most likely were. There is a correlation between college-age students and the gratification of being rewarded a pizza (Ramani 2014).

These reasons included the possible effects of control variables we were not directly interested in, yet provided evidence for these differences. In the regression analysis section, we use these controls to examine effects they may have caused on the main factors we tested.

Regression Analysis

In order to complete a more rigorous analysis of the dataset from the experiment, Dr. Bockstedt provided the skillset to write commands and interpret results from a data analysis and statistical software called STATA. Initially, we began to develop a standard regression model, but realized moving forward that the data is dichotomous – data was represented as either a 0 or 1, and so a logistic regression model was used for its ability to model binary responses. We developed a logistical regression model, used STATA to determine results using the output tables, and analyzed those results for statistical significance.

Regression Model

Our regression model is built upon the dependent and independent variables we chose from our dataset. Since we wanted to focus on sensitive information provided by subjects, we came up with an observation index around personally identifiable information (PII), which is represented as a dependent variable, DV_1. The two main factors, or predictor variables, assigned to the regression model include context, reward. Other control variables include the subject's gender, subject age, confederate gender, and subject-confederate gender match.

STATA results

Using STATA with the previous logistic regression model we developed, we utilized the logit command and entered in the related covariates. We used this command because of the nature of results in our dataset. Since the results are dichotomous (either subjects provided PII or they did not), the standard regression would not provide credible results. With this, we first looked at PII

as the dependent variable with the 5 previously discussed covariates as predictor variables and used logit to estimate our logistic regression model. The results are detailed in a table below for clarity. A screenshot of results from the STATA software can be found as an item in Appendix E.

Variable	Coefficient	Standard Error	P > |z|
High Reward (iPad)	**-1.407759**	0.6901394	**0.041**
Charity Context (BFK)	-1.003127	0.6244619	0.108
High Reward w/ Charity Context	**2.14152**	0.9608209	**0.026**
Subject Gender	0.2901283	0.4844069	0.549
Subject Age	0.018537	0.0279989	0.508
Confederate Gender	0.5411065	0.4794704	0.259
Gender Match	1.144382	0.4879876	0.019

Table 5: Logistic Regression Results for DV_1 - PII

By looking at the first variable for "High Reward (iPad)" and the associated highlighted cells in yellow, the coefficient value is approximately -1.41, with a P>|z| score of 0.041. This value means the subject was 1.41 times **less likely** to provide their PII on the survey to the confederate. The second yellow cell shows a P>|z| value of 0.041, which is less than 0.05, meaning that there is strong evidence against the null hypothesis. We hypothesized that subjects would provide more PII given a high reward (Hypothesis 1 – Main Effect), although the results show that subjects were in fact less likely to do so. This goes against our hypothesis, but otherwise is still

interesting to note. As discussed previously, there could have been numerous confounding factors as to why this is the case.

By looking at the third variable down for "High Reward w/ Charity Context" and the associated highlighted cells in green, the coefficient value is approximately 2.14, with a P>|z| score of 0.026. This item signifies the interaction between the Books-For-Kids *context* with the iPad Mini *reward*. Looking at the first green cell, this value means the subject was 2.14 times **more likely** to provide their PII on the survey to the confederate. The second green cell shows a P>|z| value of 0.026, which is less than 0.05, meaning that there is strong evidence against the null hypothesis. This result confirms our third hypothesis (Hypothesis 3 – Interaction Effect) that the combination of a high reward with a charitable organization will make subjects provide more PII. While the first hypothesis was not correct, it is interesting to note that the interaction between high reward and charity make it more likely that a subject will provide PII. The effect of going from marketing to charity increases the impacts likelihood.

With this analysis, we can conclude that a social engineer is just over two times more likely to collect personally identifiable information when posing as an individual from a charitable organization paired with a high reward. While the specific charitable organization and "high" reward used in this experiment may have an effect towards how a potential victim may react, this case shows that social engineers do not have to try very hard to collect more information from victims. In our experiment design, the raffle for rewards were, in fact, real. For a social engineer, just the *promise* of a reward through a raffle could be enough to gather sensitive information.

Discussion

Based on the analysis and results from our study, it is clear that a would-be social engineer can easily target, lure, and obtain information from unsuspecting victims by posing as an individual from a charitable organization while offering potentially high-value rewards. While it is not 100% of the time that subjects provided their personal information, an alarming amount of people provided the most sensitive information to the surveyors. With the information provided, a social engineer would be able to open up credit card accounts, reroute money, impersonate individuals, and commit other heinous fraudulent activities. As discussed previously, the emotional and financial recovery from fraud abuse is devastating and disruptive, even if insurance can save some aspect of the victim's despair.

Given that the threat landscape for practitioners and businesses is all-encompassing, including retail, government, e-commerce, financial, and so on, there is a great amount of value in the study and results we've achieved here. Social engineering is a great threat to information security and cybersecurity, although there are some interesting conclusions that can be drawn from the results. Knowing that social engineers can manipulate factors such as context and reward suggests that the managers of information and training can begin to build an awareness for these target vectors. That is, people can build threat indicators around factors regarding when social engineering may be more likely to happen. For example, a retail company selling textbooks to students can benefit from a threat indicator that is triggered when an employee is being possibly socially engineered by the use of context and reward factors. The retail company can then act on the possible attack and deny the attacker whatever information they were trying to glean from the retail company's employee.

Future research could also learn from our experiment by conducting further studies into social engineering through other factors as well. Given that our study identified some issues, while also providing results from quantitative, qualitative, and introspective views, there is a plethora of changes and opportunities for further research in social engineering. The results focused on two factors, although there are many other avenues this research can move into, such as using distraction as a factor, temporal effects on the subject for post-survey learning over time, credibility, and how confederate attire affects results, for example. Given the potential confounding factors we identified, confederates themselves could be better controlled during the surveying approach, location changes, and the sample population could be focused on a non-biased group of people. Logistically, there was much to accomplish in order to go from start to end in data collections given the personnel requirement, materials, protocols, and approvals for implementing a deceptive survey experiment.

The implications of our study show that social engineering is relatively simple to conduct and potentially highly damaging, although the results can be investigated to provide a mitigation strategy moving forward. With the academic research conducted here, society can learn and benefit from more informed training and a greater awareness of information importance and risk. Having gone through the process to conduct a field experiment through the rigors of academic research, while also learning more deeply a few of the factors of social engineering, the contribution of knowledge for understanding greater why people make decisions is complex yet it can be understood. Taking this research design and moving forward, researchers can develop other models that can shape the understanding of decision making through other studies. It is important for this research to continue, since the implications affect all levels of organizations,

their employees, as well as policy. The protection of PII is important to deter identity theft and fraud, and this research paves a way for academics to better understand how to protect individuals' PII by conducting social engineering studies, field experiments, and analysis.

Conclusion

Our research idea, paired with a unique design for social engineering, has led to a strong sample collection for the four combinations of factors that describe a subject's antecedence to social engineering susceptibility. With enough data points, we were able to identify trends and patterns in the data to determine summary statistics within and between the categories and questions answered by subjects. By applying statistical analyses, we will be able to provide insight as well as meaningful results that validated and refuted our hypotheses. The implications of our results will be able to apply towards the formation of better training in the corporate and public sectors, as well as have a direct application in generating policies. This research also allows for the ability to continue research in the field of behavioral psychology as it relates to information security and social engineering.

References

Applegate, S. (2005). Social Engineering: Hacking the Wetware!. *Information Security Journal: A Global Perspective.* Volume 18, Issue 1, 40-46.

Bureau of Labor Statistics, Department of Justice. *www.bjs.gov*. 2013. Web. 17 Dec. 2015.<http://www.bjs.gov/content/pub/pdf/vit12.pdf>.

Colin, C., & George, L. (2004). *Behavioral economics: Past, present, future*. Princeton: Princeton University Press.

Dimensional Research. The Risk of Social Engineering on Information Security: A Survey of IT Professionals. *www.greycastlesecurity.com*. September 2011.Web.09 February 2015. <https://www.greycastlesecurity.com/resources/documents/The_Risk_of_Social_Engineering_on_Information_Security_09-11.pdf>.

Finn, P., Jakobsson, M. (2007). Designing Ethical Phishing Experiments. *IEEE Technology and Society Magazine.* Issue: Spring 2007. 46-58.

Gilovich, T., Griffin, D. W., & Kahneman, D., 1934. (2002).*Heuristics and biases: The psychology of intuitive judgment*. New York;Cambridge, U.K;: Cambridge University Press.

Goodchild, J. *www.networkworld.com.* 2010. Web. 17 Dec. 2015. <http://www.networkworld.com/article/2213486/security/the-robin-sage-experiment--fake-profile-fools-security-pros.html>.

Grossklags, J., Acquisti, A. (2005). Privacy and Rationality in Individual Decision Making. *IEEE Journals & Magazines*. Volume: 3, Issue: 1, 26-33.

Grossklags, J., Acquisti, A. (2007). When 25 Cents is too much: An Experiment on Willingness-To-Sell and Willingness-To-Protect Personal Information. *Proceedings of Sixth Workshop on the Economics of Information Security (WEIS 2007).* June 7-8, 2007.

Holt, C. A., & Laury, S. K.. (2005). Risk Aversion and Incentive Effects: New Data without Order Effects. *The American Economic Review*, *95*(3), 902–904. Retrieved from http://www.jstor.org/stable/4132749

Jagatic, T., Johnson, N., Jakobsson, M., Menczer, F. (2007). Social Phishing. *Communications of the ACM.* October 2007. Vol. 50, No. 10.

Luo, X., Brody, R., Seazzu, A. (2011). Social Engineering: The Neglected Human Factor for Information Security Management. *Information Resources Management Journal.* Volume 24, Issue 3. (pp.1-8).

McDermott, B. (2016). Factors Enabling Fraud: A Study of Social Engineering and Identity Theft. *University of Arizona, Master's Report.*

Peltier, T. (2006). Social Engineering: Concepts and Solutions. *EDPACS, 33:8, 1-13*.

Petty, R., Cacioppo, J., Schumann, D. (1983).Central and Peripheral Routes to Advertising Effectiveness: The Moderating Role of Involvement. *Journal of Consumer Research*. Volume 10, Issue 2. (pp. 135-146).

Plous, S. 1993. *The Psychology of Judgment and Decision Making*. Journal of Marketing, July 1994. Vol. 58, Issue: 3, pp. 119-120.

Ramani, N. (2014). *College students like pizza… A LOT. A new study even says so.* USATODAY COLLEGE. *college.usatoday.com.* February 25, 2014. Web. 04 April 2016. < http://college.usatoday.com/2014/02/25/college-students-like-pizza-a-lot/>.

Ross, A. (2008). *Security Engineering: A Guide to Building Dependable Distributed Systems*. 2nd ed. Indianapolis, Ind: Wiley Pub, 2008. Web.

Schwarz, N. (2000). Emotion, cognition, and decision making. *Cognition & Emotion, 14*(4), 433-440. doi:10.1080/026999300402745

Slovic, P., Peters, E., Finucane, M., MacGregor, D. (2005). Affect, Risk, and Decision Making. *Health Psychology*. July 2005. Vol. 24, Issue: 4S, S35-S40.

Software Engineering Institute (SEI). Unintentional Insider Threats: Social Engineering. *sei.cmu.edu*. January 2014. Web. 09 February 2015. <https://resources.sei.cmu.edu/asset_files/TechnicalNote/2014_004_001_77459.pdf>.

Workman, M. (2007). Gaining Access with Social Engineering: An Empirical Study of the Threat. *Information Systems Security.* Volume 16, Issue 6. (pp. 315-331).

Workman, M. (2008). A test of interventions for security threats from social engineering. *Information Management & Computer Security*. Volume 16, Issue 5. (pp. 463-483).

Appendix A

Script for Confederate

"Excuse me sir/ma'am, would you like to do a survey for a chance to win a (free iPad mini / free pizza dinner)? I am working with the (Books For Kids / BNI Market Research) organization and we are doing this because…"

1. Books for Kids

"Collecting data helps us to compare the needs of demographic groups in the US Southwest. It's for a good cause."

2. BNI Market Research

"Collecting data helps us to understand the needs of various local market segments and how they might respond to certain products."

[If they accept] "Thanks, here is the survey and a pen."

[If they decline] "That's okay, you can still enter the raffle if you would like by completing a raffle form." [Provides actual raffle form.]

Appendix B

Confederate FAQ

What is BNI Market Research, Inc?

BNI is a for-profit organization that specializes in gathering data for regional segments using live survey data. Our research ranges includes automobiles, technology, media, and more.

What is BFC 501(c)(3)?

BFC is a non-profit organization aiming to reduce the illiteracy among children between 2nd and 4th grade in the southwest region of the United States.

What will you do with my information?

We compile all of the results, anonymize them, and use them to tailor products better suited to the tastes of people in the Southwestern US. Your answers help us understand what kinds of products people are really interested in.

Do I need to fill out all the questions?

Participation is strictly voluntary. If you do not fill out all of the information, you will still be entered in the raffle, but we will not be able to conduct our research as effectively.

Will you contact me with further offers?

No, we will not contact you for commercial purposes, but we will contact you in the event that you are the winner of the raffle.

Do I need to buy anything?

There is no purchase necessary to enter the raffle.

If I win, when will I receive the prize?

Please allow 4-6 weeks to receive the prize. If you are the winner, we will contact you at the phone number and/or email address that you have provided.

What if I want to know more?

After you complete the survey, our representative will be happy to explain more about our services.

Appendix C

Associate Researcher Debrief Script

Thank you for participating in this research. Please allow me to explain the context of the survey. It was actually conducted as part of a University of Arizona study on persuasion and human behavior. We will destroy any information that you have provided to us today.

> [*The confederate will make note of which questions were answered on a separate document, or possibly on a perforated section of the same document, and either return the survey to the subject or destroy it on site. We may need to have a paper shredder available under the booth*.]

Your information is truly helpful in understanding the factors related to why individuals choose or refuse to provide personal information. The raffle is real and we will contact you if you are the winner. We will not sell or distribute your email address to anyone. Have a great day.

***** If the subject does not engage, they can still take the survey.**

Script: If you don't want to take the survey, you can still enter the raffle by going to this website and submitting your email address? [Hands the subject a card with a URL link to a Google Form that we have created.]

Appendix D

Backstory for Confederates

Books for Kids Charity:

What the organization is

BFC is a non-profit organization aiming to reduce illiteracy among children in the Southwest region of the United States. By gathering donated books and distributing them to needy schools, BFC moves closer to accomplishing its goal of stronger literacy among children in 2nd-4th grade. In order to accomplish this, BFC is working with local non-profit organizations in the Tucson and Phoenix areas to locate schools in need of appropriate reading materials that can enhance the reading abilities of young children.

Why we are collecting information

We are collecting survey data for different market segments in the Southwest region to conduct an analysis to see what the differences are among the Southwest demographic areas. The data we are looking for will be compiled, analyzed, and given to the BFC researchers for evaluation. From there, we can optimize our distribution of books for children with the greatest need.

What the information will be used for

The questions listed on the survey have been researched to provide quality results in generalizing demographics about specific areas. In order to best fulfill the BFC's goal of identifying which schools/areas are in most need, we are requesting the responses to these questions. If you have any more questions, please refer to our Frequently Asked Questions (FAQ) document.

BNI Market Research:

What the organization is

BNI is a for-profit organization that specializes in gathering data for regional segments using live survey data. Our research ranges includes automobiles, technology, media, and more. By gathering market data, BNI moves closer to accomplishing its goal of providing organizations with better business intelligence, which translates to better customer targeting campaigns.

Why we are collecting information

In order to better understand how to target specific customers for certain products, we are requesting responses to these questions in order to conduct an analysis to see the differences between southwest demographic areas. BNI conducts the market research via surveys for organizations to identify key market areas and their interests.

What the information will be used for

We will use this information to generate a map with clusters of groups with generalized interests. This information will allow organizations who are requesting the information to provide more targeted marketing solutions.

Appendix E

STATA Logistic Regression Results

```
. logit PII i.charity##i.highreward female age confedsex gendermatch if confederatejastif != "TIFF"

Iteration 0:   log likelihood = -65.395132
Iteration 1:   log likelihood = -59.395172
Iteration 2:   log likelihood = -59.199178
Iteration 3:   log likelihood = -59.198533
Iteration 4:   log likelihood = -59.198533

Logistic regression                             Number of obs     =        113
                                                LR chi2(7)        =      12.39
                                                Prob > chi2       =     0.0883
Log likelihood = -59.198533                     Pseudo R2         =     0.0948

------------------------------------------------------------------------------
         PII |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    1.charity |  -1.003127   .6244619    -1.61   0.108    -2.22705     .220796
1.highreward |  -1.407759   .6901394    -2.04   0.041    -2.760407   -.0551103
             |
charity#high~d |
         1 1 |   2.14152   .9608209     2.23   0.026     .2583458    4.024694
             |
      female |   .2901283   .4844069     0.60   0.549    -.6592917    1.239548
         age |    .018537   .0279989     0.66   0.508    -.0363399    .0734139
   confedsex |   .5411065   .4794704     1.13   0.259    -.3986382    1.480851
gendermatch |   1.144382   .4879876     2.35   0.019     .1879443     2.10082
       _cons |  -1.753347   .8333073    -2.10   0.035      -3.3866   -.120095
------------------------------------------------------------------------------
```

Note: the last section of the command included an exclusion of "TIFF". Since one of the confederates (Tiffany = TIFF) did not provide a legitimate experience throughout the experiment, her data and results have been removed from the analysis.