Software Engineering Institute

Unintentional Insider Threats: Social Engineering

The CERT® Insider Threat Center

Produced for
Department of Homeland Security
Federal Network Resilience Cybersecurity Assurance Branch

January 2014

TECHNICAL NOTE
CMU/SEI-2013-TN-024

CERT® Division

http://www.sei.cmu.edu

Carnegie Mellon University

This material is based upon work funded and supported by Department of Homeland Security under Contract No. FA8721-05-C-0003 with Carnegie Mellon University for the operation of the Software Engineering Institute, a federally funded research and development center sponsored by the United States Department of Defense.

This report was prepared for the
SEI Administrative Agent
AFLCMC/PZM
20 Schilling Circle, Bldg 1305, 3rd floor
Hanscom AFB, MA 01731-2125

DM-0000579

Table of Contents

Sponsor Information vii

Acknowledgments ix

Executive Summary xi

Abstract xv

1 Introduction 1

2 Background 3
- 2.1 Contributing Factors Described in Initial Unintentional Insider Threat (UIT) Study 3
- 2.2 Feature Model Developed in Initial UIT Study 4
- 2.3 Findings and Recommendations of Initial UIT Study 5

3 Defining and Characterizing UIT 6
- 3.1 Definition of UIT 6
- 3.2 Definition of Social Engineering 6
- 3.3 Social Engineering Taxonomy 7

4 Review of Research on Social Engineering UIT Incidents 11
- 4.1 Research on Demographic Factors 13
  - 4.1.1 Gender 13
  - 4.1.2 Age 13
  - 4.1.3 Personality Traits 14
  - 4.1.4 Culture 16
  - 4.1.5 Summary 16
- 4.2 Research on Organizational Factors 17
  - 4.2.1 Inadequate Management and Management Systems 17
  - 4.2.2 Insufficient Security Systems, Policies, and Practices 18
  - 4.2.3 Job Pressure 19
  - 4.2.4 Summary 19
- 4.3 Research on Human Factors 20
  - 4.3.1 Lack of Attention 20
  - 4.3.2 Lack of Knowledge and Memory Failure 21
  - 4.3.3 Faulty Reasoning or Judgment 21
  - 4.3.4 Risk Tolerance and Poor Risk Perception 22
  - 4.3.5 Casual Values and Attitudes About Compliance 22
  - 4.3.6 Stress and Anxiety 23
  - 4.3.7 Physical Impairment 23
  - 4.3.8 Summary 24

5 Summary of Collected Cases 26
- 5.1 Representative Cases 26
  - 5.1.1 Single-Stage Phishing Attacks 27
  - 5.1.2 Multiple-Stage Phishing Attacks 29
- 5.2 Characterization of Case Study Data 30
  - 5.2.1 Demographic, Organizational, and Human Factors 30
  - 5.2.2 Discussion and Implications of Sample Data Obtained to Date 31

6 Conceptual Models for Social Engineering Incidents 33
- 6.1 Attack Progression Analysis 33

6.2 Patterns Inferred from UIT Case Studies 34
6.2.1 Single-Stage Phishing Attack 35
6.2.2 Multiple-Stage Phishing Attack 38
6.3 Descriptive System Dynamics Model 45
6.3.1 Causal Loop Diagrams 45
6.3.2 Confirmatory Bias Loop 46
6.3.3 Phishing Exploits in Social Engineering 46
6.3.4 Confirmatory Bias in Social Engineering 47
6.3.5 Integrated Model of the Social Engineering Problem 48
6.4 Ontology of Social Engineering Tactics 50
6.4.1 Need for a Taxonomy 50
6.4.2 Social Engineering Tactics Described in Research Literature 50
6.4.3 Design Goals for the Taxonomy 51
6.4.4 The Taxonomy 52
6.5 Implications for Mitigation of Social Engineering Exploits 53
6.5.1 Implications of Patterns and Characterizations 54
6.5.2 Implications of Social Engineering Tactics Ontology 55
6.5.3 Implications of System Dynamics Model 56
6.5.4 Summary and Conclusions About Mitigation 59

7 Conclusions 60
7.1 Overview of Findings 60
7.2 Research Needs 61
7.2.1 Assessment of State of Practice and Effectiveness of Tools 61
7.2.2 Development of an Extensive UIT Database 61
7.2.3 Detailed Analysis of UIT Incidents 63

8 Recommendations 64

Appendix A: Possible Contributing Factors in Social Engineering Susceptibility 67

Appendix B: Case Study Material 71

References 82

List of Figures

Figure 1:	Social Engineering Taxonomy (Branch of Interest to This Study Highlighted)	8
Figure 2:	Single-Stage Phishing Attack, Example 1.	28
Figure 3:	Single-Stage Phishing Attack, Example 2.	28
Figure 4:	Single-Stage Phishing Attack, Example 3.	29
Figure 5:	Multiple-Stage Phishing Attack, Example 1.	29
Figure 6:	Cloppert's Six-Phase Attack Progression	33
Figure 7:	Workflow Pattern Showing Phases of a Single-Stage Phishing Attack	35
Figure 8:	Use Case Model for Single-Stage Social Engineering Attack	36
Figure 9:	Attack Class Model for a Social Engineering Attack	36
Figure 10:	Swim-Lane Chart of Actions Taken by Attacker and UIT Victims in a Single-Stage Attack	37
Figure 11:	Interaction View Showing Object Collaboration in a Single-Stage Social Engineering Attack	38
Figure 12:	Workflow Diagram Attack Chain for Multiple-Stage Phishing Exploit	39
Figure 13:	Use Case Model of a Multiple-Stage Social Engineering Attack	41
Figure 14:	Interaction View Showing Object Collaboration in a Multiple-Stage Social Engineering Attack	42
Figure 15:	Illustration of Concepts and Patterns Applied to Case #15	43
Figure 16:	Illustration of Concepts and Patterns Applied to Case #5	44
Figure 17:	System Dynamics Notation Used in Abstract Models	45
Figure 18:	Confirmatory Bias	46
Figure 19:	Causal Loop Diagram of Phishing Exploits	47
Figure 20:	Confirmatory Bias in Social Engineering Exploits	48
Figure 21:	Causal Loop Diagram of Social Engineering of Insiders by Outsiders	49
Figure 22:	Mitigation Strategies that Apply to Different Phases of an Attack	55
Figure 23:	Causal Loop Diagram of Avenues for Social Engineering Mitigation	58

List of Tables

Table 1:	Summary of Social Engineering Characteristics	10
Table 2:	Social Engineering Factors Studied by Workman	16
Table 3:	Steps in a Single-Stage Phishing Attack	35
Table 4:	Steps in a Multiple-Stage Phishing Attack	40
Table 5:	Social Engineering Tactics, Vulnerabilities, and Mitigations	56
Table 6:	Summary of Research Findings	67

Sponsor Information

The Department of Homeland Security (DHS) Office of Federal Network Resilience (FNR) Cybersecurity Assurance Branch (CAB) sponsored this report. The key contact is Project Lead Sean McAfee (sean.mcafee@hq.dhs.gov). Please forward any questions about this work to FNR/CAB via Mr. McAfee.

Acknowledgments

The Insider Threat team would like to thank John Bergey, Sholom Cohen, Matthew Collins, Jennifer Crowley, Frank Greitzer, David Mundie, Arley Schenker, and Kahlil Wallace for their integral contributions to this report.

Executive Summary

Insider threat is recognized as a major security risk by computer and organizational security professionals, more than 40% of whom report that their greatest security concern is employees accidentally jeopardizing security through data leaks or similar errors.[1] A previous report by the CERT® Insider Threat team, part of Carnegie Mellon University's Software Engineering Institute, provided an initial examination of the unintentional insider threat (UIT) problem, including an operational definition of UIT, a review of relevant research on possible causes and contributing factors, and a report on frequencies of UIT occurrences across several categories.[2] This initial work served to inform government and industry stakeholders about the problem and its potential causes and to guide research and development (R&D) investments toward the highest priority R&D requirements for countering UIT. The current effort seeks to advance our understanding of contributing factors to UIT by focusing on UIT incidents involving social engineering. The goals of this project are to collect additional UIT incident data to build a set of social engineering cases to be added to the CERT Division's Management and Education of the Risk of Insider Threat (MERIT) database (referred to as the *insider threat database*), and to analyze UIT cases to identify possible behavioral and technical patterns and precursors, with a particular focus on social engineering cases. We hope that this research will inform future research and development of UIT mitigation strategies.

Defining and Characterizing Unintentional Insider Threat (UIT)

Based on our original UIT study and the current study, we define *UIT* as the following:

> *An unintentional insider threat is (1) a current or former employee, contractor, or business partner (2) who has or had authorized access to an organization's network, system, or data and who, (3) through action or inaction without malicious intent, (4) unwittingly causes harm or substantially increases the probability of future serious harm to the confidentiality, integrity, or availability of the organization's resources or assets, including information, information systems, or financial systems.*

We define social engineering, in the context of UIT incidents, as the following:

> *Social engineering, in the context of information security, is manipulation of people to get them to unwittingly perform actions that cause harm (or increase the probability of causing future harm) to the confidentiality, integrity, or availability of the organization's resources or assets, including information, information systems, or financial systems.*

We created a preliminary social engineering taxonomy that is consistent with descriptions of social engineering exploits in the scientific literature as well as real cases reported in court documents and other print media. This taxonomy reinforces the definition provided above and

---

1 AlgoSec. *The State of Network Security 2013: Attitudes and Opinions.* AlgoSec, Inc., 2013. http://www.algosec.com/resources/files/Specials/Survey%20files/State%20of%20Network%20Security%202013_Final%20Report.pdf

® CERT® is a registered mark owned by Carnegie Mellon University.

2 Insider Threat Team, CERT. *Unintentional Insider Threats: A Foundational Study* (CMU/SEI-2013-TN-022). Software Engineering Institute, Carnegie Mellon University, 2013. http://www.sei.cmu.edu/library/abstracts/reports/13tn022.cfm

provides a mutually exclusive, exhaustive organization of the various forms of social engineering exploits. Our research focuses on the portion of the taxonomy that applies to UIT incidents.

Research Findings on Possible UIT Contributing Factors

In the initial phase of UIT research already reported to DHS, we identified potential causal and correlational factors for all UIT cases in our database; some of those factors are also relevant to social engineering exploits. We use the term *correlational factors* because the causal relationship between each of these factors and the frequency of social engineering has not been empirically identified. The initial phase of research identified the factors through a review of literature in the related fields of human factors, cognitive psychology, human error, and decision making, and the factors were then organized into several broad categories.

Our current research effort, focused on UIT social engineering exploits (such as phishing), sorted the initial set of possible contributing factors into three categories: demographic, organizational, and human factors. Relevant research and case study data informed our conceptual modeling efforts to characterize UIT social engineering exploits.

Regarding possible demographic factors, there is limited support for the notion of individual differences in phishing susceptibility across the demographic factors of age, gender, and personality; not enough research is available to determine possible cultural differences. Because relatively few publications exist on the possible contributions of demographic, organizational, and human factors to UIT social engineering susceptibility, our review included the human error literature; in some cases, a UIT incident can be attributed to human error.

Some organizational factors can increase the likelihood of human errors (i.e., lapses in judgment) at the employee level (the following list is not exhaustive): (a) poor management or management systems that may fail to assign sufficiently qualified personnel to tasks or that provide employees insufficient materials and resources, (b) inadequate information security systems or policies, and (c) work environments or work planning and control systems that impact employee satisfaction or cause stress or anxiety. Many human factors variables have also been identified as more immediate causal factors: lack of attention or lack of knowledge, which often cause people to ignore security cues, and a tendency to focus disproportionately on urgency cues. Phishers exploit these cognitive limitations by employing visual deception to spoof legitimate email messages or websites and by appealing to the victim's willingness to help in urgent situations.

Susceptibility to social engineering attacks also may be traced to the tendency for individuals to underestimate and ignore the threats, particularly under conditions of high workload. Risk tolerance and perception represents another significant human factor: individuals who are less risk-averse are more likely to fall for phishing schemes. In addition, individuals might ignore these threats because they perceive information security compliance as interfering with job functions.

To the extent that relevant data may be obtained or inferred from reports of UIT incidents, these demographic, organizational, and human factors should be tracked and maintained in a UIT incident database. Analysis of trends will serve to inform and prioritize the development of enterprise-level mitigation strategies.

Case Studies, Characteristics, and Patterns of Social Engineering Exploits

This report describes some UIT case studies involving social engineering exploits; Appendix B summarizes the cases used in our analysis. We apply analytical methods to gain a better understanding of the problem, highlight common features across multiple exploits, and identify possible mitigation strategies. These analytical methods include attack progression analysis, characterization of attack patterns, system dynamics modeling, and the creation of an ontology of social engineering tactics. Our analysis seeks to synthesize research and case studies to identify possible contributing factors and patterns that may be useful in designing mitigation strategies. Systematic examination of the resulting patterns and models informs concepts for mitigation approaches that may be applied to particular patterns or stages in UIT social engineering attacks.

Conclusions

There is at best a weak association between social engineering susceptibility and various demographic factors (age, gender, etc.), emphasizing the need for more research to clarify or disambiguate certain relationships. Research suggests it may be possible to use personality factors to identify individuals who are at higher risk of falling victim to social engineering or to better tailor training topics for vulnerable personality factors; however, further research is necessary.

Organizational factors can produce system vulnerabilities that adversaries may exploit in social engineering attacks. Management systems or practices that provide insufficient training, inadequate security systems and procedures, or insufficient resources to successfully complete tasks may promote confusion, reduce understanding, and increase employee stress, all of which increase the likelihood of errors or lapses in judgment that enable the attacker to successfully breach defenses.

Academic research has identified human factors that may underlie UIT social engineering susceptibility, but the lack of reporting on relevant human factors in real-world cases has hampered validation of potential human factors. Academic research suggests that relevant human factors include insufficient attention or knowledge that would enable users to recognize cues in socially engineered messages; cognitive biases or information processing limitations that may lead the UIT victim to succumb to deceptive practices and obfuscation; and attitudes that ignore or discount risks, or that lead individuals to take shortcuts around information-security compliance they feel is interfering with job functions.

Analysis and conceptual modeling of collected case studies reveal a number of commonalities or patterns that may inform the development of mitigation tools or strategies. Social engineering attacks may be characterized as comprising a single stage or multiple stages, and within each stage there are recognizable patterns or building blocks that compose the attack.

These conclusions suggest several research needs, including further study of organizational and human factors as well as additional case study data. To advance the current practice and state of the art in computer and network defense, and especially safeguards against social engineering, the following research needs should be addressed:

- Assess state of practice and effectiveness of mitigation tools and approaches.
- Develop an extensive, self-reporting UIT database.
- Conduct detailed analysis of UIT social engineering incidents to inform development of more effective mitigation approaches and tools.

Recommendations

The research community as well as responsible organizations and stakeholders are obligated to continue research and information gathering to inform the development of effective training and mitigation tools. Our review and analysis of research and case studies suggests the following strategies to reduce the effectiveness of social engineering attacks.

1. Continue to record demographic information as case studies are tabulated and entered into the UIT database. The records should include the demographic factors described in this report.
2. Organizations should ensure that their management practices meet human factors standards that foster effective work environments to minimize stress (e.g., minimizing time pressure and optimizing workload) and encourage a healthy security culture.
3. Organizations should develop and deploy effective staff training and awareness programs aimed at educating users about social engineering scams, including learning objectives to help staff attend to phishing cues, identify deceptive practices, and recognize suspicious patterns of social engineering exploits. Training objectives should also include effective coping and incident management behaviors (ways to overcome one's own limitations and susceptibilities as well as appropriate responses to social engineering exploits).
4. The research and stakeholder community should develop mitigations that apply to specific attack phases as described in this report (i.e., research and open source intelligence phase, planning and preparation phase, launch operation phase, information capture phase, and culmination/exploitation phase).

Countering the UIT social engineering problem poses major challenges to organizations, who must balance operational goals with security goals to maintain a competitive edge in the market. Because organizational policies and practices are resistant to change, it is a great challenge to keep up with the rapidly changing, increasingly sophisticated social engineering attacks. Some social engineering campaigns may be so well crafted that they can defeat the organization's best countermeasures (e.g., training and policies). An attack can succeed if only one employee succumbs to an exploit, so an organization's strategy to combat UIT social engineering must be comprehensive and include cybersecurity tools, security practices, and training. By characterizing and conceptually modeling the UIT social engineering problem, this report has sought to inform mitigation development efforts and identify research needs to more effectively combat UIT social engineering exploits.

Abstract

The research documented in this report seeks to advance the understanding of the unintentional insider threat (UIT) that derives from social engineering. The goals of this research are to collect data on additional UIT social engineering incidents to build a set of cases for the Management and Education of the Risk of Insider Threat (MERIT) database and to analyze such cases to identify possible behavioral and technical patterns and precursors. The authors hope that this research will inform future research and development of UIT mitigation strategies.

1 Introduction

A significant proportion of computer and organizational security professionals believe insider threat is the greatest risk to their enterprise, and more than 40% report that their greatest security concern is employees accidentally jeopardizing security through data leaks or similar errors [AlgoSec 2013]. A previous report by the CERT® Insider Threat team, part of Carnegie Mellon University's Software Engineering Institute, provided an initial examination of this problem [CERT 2013]. That report characterized the unintentional insider threat (UIT) by developing an operational definition, reviewing relevant research to gain a better understanding of its possible causes and contributing factors,[1] and providing examples of UIT cases and the frequencies of UIT occurrences across several categories. The report also documented our first design of a UIT feature model, which captures important elements of UIT incidents.

One challenge in researching the UIT problem and developing effective mitigation strategies is that the UIT topic has gone largely unrecognized in scientific research, and UIT incidents and case studies have gone mostly unreported. In particular, incident reports typically lack sufficient detail to inform analyses of potential contributing factors. The initial work of the CERT Insider Threat team [CERT 2013] served to inform government and industry stakeholders about the problem and its potential causes and to guide research and development (R&D) investments toward the highest priority R&D requirements for countering UIT. As a follow-on to that study, the current effort sought to advance our understanding of UIT contributing factors by focusing on a major type of UIT incident, social engineering. The goals of this research project were to

- collect additional UIT incident data to build a set of social engineering cases to be added to the CERT Division's Management and Education of the Risk of Insider Threat (MERIT) database (referred to as the *insider threat database*), which documents insider threat cases
- analyze UIT cases to identify possible behavioral and technical patterns and precursors, with a particular focus on UIT cases that involve social engineering, to inform future research and development of UIT mitigation strategies

This report documents progress in meeting these objectives. The remainder of the report is organized as follows:

- Section 2, Background, provides a brief overview of work accomplished in the initial phase of work.[2]
- Section 3, Defining and Characterizing UIT, updates the definitions of UIT and, in particular, social engineering exploits, which are the main focus of this study.
- Section 4, Review of Research on Social Engineering UIT Incidents, updates our literature review, focusing on social engineering UIT incidents.

® CERT® is a registered mark owned by Carnegie Mellon University.

1 A *factor* is a situational element or feature that may or may not be related to the existence of the incident. A *contributing factor* is a factor that has been demonstrated to be associated as a causal factor of an incident. Because in general causal relationships have not been shown, our usage of the term *contributing factor* should be interpreted as *potential contributing factor*.

2 The initial phase of work is reported in *Unintentional Insider Threats: A Foundational Study* [CERT 2013].

- Section 5, Summary of Collected Cases, describes the case collection requirements we developed to guide collection and reporting of UIT cases. This section also provides examples of representative UIT cases involving social engineering exploits.
- Section 6, Conceptual Models for Social Engineering Incidents, discusses results synthesized from our research and case study analyses to identify patterns that may be useful in designing mitigation strategies.
- Sections 7 and 8 discuss conclusions and recommendations, respectively.
- Appendix A provides additional details on contributing factors.
- Appendix B provides additional details on case study data.

2 Background

2.1 Contributing Factors Described in Initial Unintentional Insider Threat (UIT) Study

In our initial phase of work [CERT 2013], we observed that the applicable research on UIT may be organized in several different ways. One useful way is to identify potential causal and correlational factors of UIT incidents, including those pre-existing factors that may increase the likelihood of a UIT incident, the series of events leading up to the attack, and the steps involved in the attack itself. Thus, this report is loosely organized into sections based on potential factors identified and the steps involved in the attack, which we call an *attack pattern*. This phase of the research effort focuses exclusively on a class of UIT threats involving social engineering, which is formally defined in Section 3.1.

Part of the UIT definition we generated in the initial phase includes humans' failure to appropriately identify and respond to UIT threats, which can be partially attributed to human cognitive limitations and biases in perception and decision making. While the adversary will typically penetrate a network at the individual employee level, we recognize that factors inherent in the context of the UIT incident can contribute to the employee's vulnerability. Human errors, in the context of UIT, may never be eliminated completely, but human error mitigation techniques may dramatically reduce errors that allow adversaries to penetrate the network. Research in workplace safety and ergonomics suggests that mitigation strategies should include the identification of these contextual factors (e.g., organizational practices and policies, adversarial sophistication) that contributed to errors and resultant adverse outcomes [Dekker 2002, Pond 2003].[3]

In our initial-phase review of literature in the related fields of human factors, cognitive psychology, human error, and decision making (but not specifically addressing UIT) [CERT 2013], we identified a large set of *possible* contributing factors to UIT incidents and organized these factors into several broad categories, following Pond's work aimed at identifying factors leading to security incidents [Pond 2003]. We described deep-seated *organizational factors* in terms of problems with data flow, work setting, work planning and control, and employee readiness: These organizational factors may raise the likelihood of human errors and conditions that may underlie many UIT incidents. We described more immediate correlates of UIT incidents in terms of a diverse set of *human factors*, including lack of situation awareness, issues relating to risk tolerance and risk perception, inadequate knowledge, flawed reasoning and decision making, and illness, injury, and other health-related factors that diminish decision making, judgment, or other cognitive capabilities. We also considered possible associations with *demographic factors* such as age, gender, and cultural factors. This research took initial steps toward (a) describing or speculating on possible mechanisms by which these diverse factors might influence the occurrence of UIT incidents and (b) examining case studies to determine which, if any, of these factors have been documented as possible contributing causes in published UIT cases. This initial

[3] An important difference between general human error and human error in social engineering incidents is that the latter involves a malicious adversary who employs considerable obfuscation techniques to fool the unwary victim. A sophisticated social engineering attack may well succeed despite the organization's mitigation strategies (policies, tools, training, management practices, etc.).

work did not specifically focus on factors relating to UIT exploits involving social engineering. A conventional approach to human error analysis does not typically account for an active adversary, so it may not fully address the underlying causal factors for social engineering exploits.

As we discuss in Section 4, our current efforts to investigate and synthesize these potential factors further has led to a more parsimonious list of possible contributing factors to focus our research, facilitate model development, and inform case data collection.

2.2 Feature Model Developed in Initial UIT Study

A *feature model* is the collection of features that characterize instances of a concept. The initial phase of work developed a feature model of a UIT incident [CERT 2013]. The model represents relevant characteristics of any UIT incident and comprises a hierarchical diagram that decomposes the concept into features and subfeatures, definitions of each feature, rules for combining features such as features requisite for other features, and rationale for choice of features. The model categorizes four mandatory features for each incident:

- roles of the individuals in a UIT incident
- possible underlying causes, correlations, and contributing factors
- system information and the format of the disclosed data
- industry sector or government agency where the incident occurred

We use the feature model to categorize cases collected and determine how frequently cases in each category occur. The analysis first considers the occurrence frequency of types of incidents under each top-level feature and its immediate subordinate features. The feature model also helps characterize threat vectors and basic patterns of activity for each incident category, allowing our researchers to use features to search for specific types of incidents.

In the initial phase of work, we used the term UIT *threat vectors*[4] to refer to different types of UIT incidents that account for virtually all the incidents we collected:

- DISC, or accidental disclosure (e.g., via the internet)—sensitive information posted publicly on a website, mishandled, or sent to the wrong party via email, fax, or mail
- UIT-HACK, or malicious code (UIT-HACKing, malware/spyware)—an outsider's electronic entry acquired through social engineering (e.g., phishing email attack, planted or unauthorized USB drive) and carried out via software, such as malware and spyware
- PHYS, or improper/accidental disposal of physical records—lost, discarded, or stolen non-electronic records, such as paper documents
- PORT, or portable equipment no longer in possession—lost, discarded, or stolen data storage device, such as a laptop, personal digital assistant (PDA), smartphone, portable memory device, CD, hard drive, or data tape

Results obtained in the initial phase of work were limited due to the paucity of data collected. Generally, 49% of the UIT cases were associated with the DISC threat vector, 6% with PHYS, 28% with PORT, and 17% with UIT-HACK. With nearly half of the incidents falling in the DISC

[4] We use the term *threat vector*, instead of the more typical term *attack vector*, in the present context because the word *attack* connotes malicious intent, which is absent in unintentional acts.

category, the study determined that release through the internet and email accounted for 23% and 20%, respectively, of all UIT cases. The combined incidence rate for PHYS and PORT vectors (related to loss of electronic devices or non-electronic records) accounted for roughly one-third of the incidents. While these findings were preliminary due to the small sample size of 35 incidents, the results led to our current focus on social engineering exploits, which account for a substantial percentage of cases collected to date.

The collection of additional UIT cases and subsequent analyses of the data will improve our understanding of similarities and differences among UIT incidents based on the model's features. The accumulation and analysis of incident statistics will also ultimately help stakeholders prioritize different types of UIT threats and associated mitigation strategies with respect to organizational risk. This prioritization also informs decision makers about where and how to invest R&D money to derive the greatest protections against UIT cases.

2.3 Findings and Recommendations of Initial UIT Study

Our initial study of the UIT problem identified many possible contributing factors of UIT incidents. As we show in Section 4, in the current phase we have simplified the original list of contributing factors to facilitate data collection, analysis, and synthesis. The preliminary study also provided numerous suggestions for possible mitigation strategies. Because of the possible role of human error in UIT incidents, we recommended that countermeasures and mitigations include strategies for improving and maintaining productive work environments, healthy security cultures, and human factors that increase usability and security of systems and decrease the likelihood of human errors. Also recommended were training and awareness programs that focus on enhancing staff recognition of the UIT problem and that help individuals identify possible cognitive biases and limitations that might put them at a higher risk of committing such errors or judgment lapses. However, training and awareness programs have their limits, and human factors or organizational systems cannot completely eliminate human errors associated with risk perception and other cognitive and decision processes. Therefore, we recommended a comprehensive mitigation strategy that includes new and more effective automated safeguards that seek to provide fail-safe measures against these failures.

It is important to reiterate that we derived the set of possible UIT contributing factors identified in the Phase 1 effort from research studies in broad areas of human factors and cognitive psychology, without the benefit of studies specifically addressing UIT. Indeed, we identified the need to continue to update the database to accommodate UIT cases, to collect UIT incident data and build up a large set of UIT cases, and to conduct more focused research on factors that contribute to UIT. These recommendations have informed the approach and objectives of the present phase of research, which focuses on UIT cases that include social engineering exploits. Because many UIT incidents involving social engineering include an element of deception, it is particularly important to re-assess the possible UIT contributing factors.

3 Defining and Characterizing UIT

3.1 Definition of UIT

Our initial research produced a working definition of an unintentional insider threat: "An unintentional insider threat is (1) a current or former employee, contractor, or business partner (2) who has or had authorized access to an organization's network, system, or data and who, (3) through action or inaction without malicious intent,[5] (4) causes harm or substantially increases the probability of future serious harm to the confidentiality, integrity, or availability of the organization's information or information systems" [CERT 2013]. While collecting and analyzing UIT cases in our current effort, we recognized a need to modify the original definition slightly. One change is to emphasize that the unintentional insider's actions occur largely without the insider's knowledge or understanding of their impact; we added the term "*unwittingly*"[6] to the fourth part of the definition. A second change is to modify the description of the target of the attack to include assets other than the organization's information system such as personnel and financial systems. The revised definition is as follows:

> *An unintentional insider threat is (1) a current or former employee, contractor, or business partner (2) who has or had authorized access to an organization's network, system, or data and who, (3) through action or inaction without malicious intent, (4) unwittingly causes harm or substantially increases the probability of future serious harm to the confidentiality, integrity, or availability of the organization's resources or assets, including information, information systems, or financial systems.*

The remainder of this report is devoted to UIT threats that include a social engineering component. Most threats of this type fall into the DISC and UIT-HACK threat vectors.

3.2 Definition of Social Engineering

A *UIT incident* typically results from actions (or a lack of action) by a nonmalicious insider (although not all such cases are characterized as completely nonmalicious and individuals involved may not always be identified). The unintentional insider's actions are often in response to an attacker's social engineering activities.

We have adopted the following working definition of social engineering and related exploits, in the context of UIT incidents:

> *Social engineering, in the context of information security, is manipulation of people to get them to unwittingly perform actions that cause harm (or increase the probability of causing future harm) to the confidentiality, integrity, or availability of the organization's resources or assets, including information, information systems, or financial systems.*

---

[5] Malicious intent includes the intention to cause harm. Harm can also be caused by those who have no malicious intent (i.e., are nonmalicious), either by action or inaction, even if they knowingly break a rule (i.e., the guard who does not check badges does not mean to allow a malicious actor into the building, but he lets someone in who sets the building on fire).

[6] This definition uses the perspective of the unintentional insider, which differs from the broader definition of social engineering acts that includes the (malicious and/or intentional) perpetrator's perspective.

Social engineering represents a type of confidence scheme aimed at gathering information, committing fraud, or gaining computer system access. Social engineering, almost by definition, capitalizes on human psychology, such as cognitive limitations and biases, which attackers exploit to deceive the victim. This differs from other types of UIT incidents examined in our initial report, such as DISC cases in which an individual inadvertently discloses sensitive information without any interaction with an outside party (e.g., posting information on public databases or losing information by discarding it without destroying it). The adversary (or adversaries) masterminding the social engineering UIT incidents may have one or more malicious objectives that correspond to the intended *impact* to the organization, such as financial loss, disruption, or information compromise.

This type of exploit does not typically constitute a single attack, but rather a step that occurs within a more complex sequence of actions that compose a larger fraud scheme. We have found it useful to identify two levels of social engineering incidents:

1. *single-stage attack*—As the name implies, the exploit is carried out in a single social engineering incident. The attacker obtains information as a result of the exploit and uses this information to cause further harm to the insider's organization. The attacker does not use the information to conduct further social engineering exploits.
2. *multiple-stage attack*—The attacker capitalizes on information gained from an initial exploit to execute one or more additional social engineering exploits. Some multiple-stage exploits play out over a matter of minutes or hours, while others may last for weeks or longer as the attacker applies the compromised information to cause harm.

3.3 Social Engineering Taxonomy

Several researchers have tried varied approaches to categorizing types of social engineering attacks. For example, Peltier breaks down social engineering into two main categories: human based and technology based [Peltier 2006]. Another decomposition uses the categories of close access (essentially human-to-human), online, and intelligence gathering [Laribee 2006a]. Some combination of each of these perspectives applies: Social engineering often occurs in multiple stages, so a UIT social engineering incident may fall into multiple social engineering taxonomic categories. We have adopted a simple yet comprehensive categorization as shown in Figure 1.

At the highest level of the taxonomy, we distinguish between whether or not exploits use interpersonal interaction. While social engineering is typically thought of as an interaction between people, UIT exploits commonly begin with the attacker gathering intelligence on the individual or organization being targeted for an attack. Because this activity does not involve manipulation of a person, some analysts do not consider it to be a form of social engineering. We include it because discussions of social engineering typically do so, and because it is often performed in conjunction with a social engineering incident. One type of intelligence gathering is referred to as *dumpster diving* or *trashing* [Laribee 2006a], in which an attacker searches for sensitive information in the garbage (e.g., bank statements, pre-approved credit cards and student loan documents that are carelessly thrown away). A second type of intelligence gathering is open source research [Laribee 2006a], that includes searching websites (e.g., Facebook, company websites) for information on targets that may be exploited in a second phase of a social engineering attack.

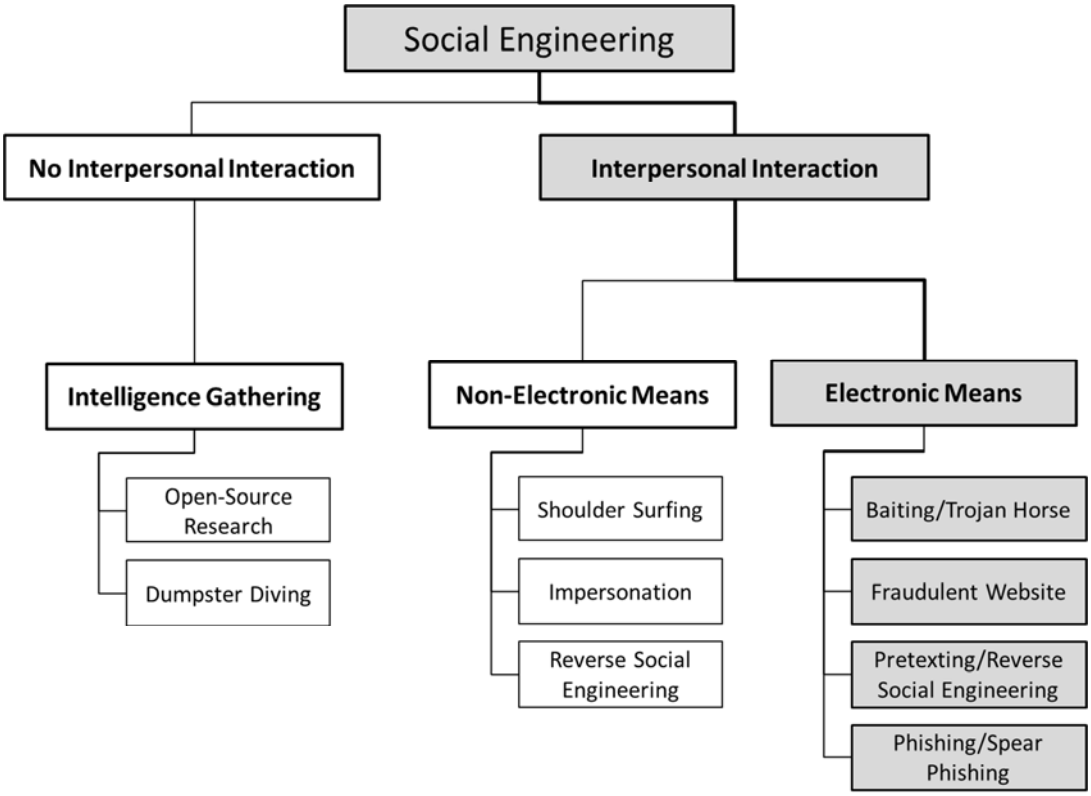


Figure 1: Social Engineering Taxonomy (Branch of Interest to This Study Highlighted)

Social engineering attacks that include interpersonal interaction may be characterized as communication directly between people (such as in person or by telephone) or mediated through electronic means (e.g., electronic media, email, and internet). We distinguish between interpersonal interaction that uses electronic means (online, using electronic media—generally, using technology) and non-electronic interaction (telephone, direct face-to-face interaction—generally, not using technology). Regardless of which type of interaction applies, these attacks are characterized by exploitation of human psychology to deceive the victims and achieve some objective (financial, sabotage, etc.).

Current literature discusses many types of non-electronic social engineering exploits. Such personal or face-to-face exploits are close-access techniques designed to gain physical access to computer systems or the information they contain. Using their people skills, social engineers use various techniques such as friendliness, impersonation, conformity, decoying, sympathy, and reverse social engineering to exploit trust relationships and gain desired information [Laribee 2006a]. One form of non-electronic social engineering is *shoulder surfing*, or stealthily looking over the shoulder of someone who enters security codes or passwords. Another broad method is *impersonation*, or creating a character and playing out a role to deceive others. Social engineering by telephone is an example of an impersonation technique, so it is not specifically shown in Figure 1. Whether by telephone or in person, an attacker who uses impersonation typically pretends to be someone in a position of authority, such as a phone company representative, bank representative, or technical support expert; the attacker calls or physically approaches a victim and

attempts to persuade the victim to provide sensitive information. A closely related exploit (not shown in Figure 1 because of its similarity to impersonation) is *tailgating*, in which the attacker poses as an employee to slip into a restricted area simply by walking behind a person with legitimate access. *Reverse social engineering* is a more sophisticated form of non-electronic social engineering,[7] in which the attacker creates a situation where the unwitting victim believes that the attacker can help solve a problem. Typically the attacker poses as a technical aide to fix a problem that the attacker created or that does not exist. The attacker communicates his or her capability to help, such as through advertising or a phone call. Finally the victim invites the attacker to assist, which eventually allows the attacker to access to the desired information.

The methods of most concern for this phase of our research are those in the Electronic Means branch of the taxonomy. Current literature describes many of these types of exploits. As was the case with the non-electronic social engineering exploits, our review of electronic social engineering exploits share many commonalities. As a result, we decided that the discussion would be simplified, without losing generality, by distinguishing the following list of representative electronic social engineering exploits (also shown in Figure 1):

- **baiting/Trojan horse**—an exploit that uses malware-infected physical media (e.g., CD-ROM, USB drive) to perpetuate an attack. The Trojan horse media look legitimate and rely on the curiosity or greed of the victim who finds the device and uses it. Insertion of the device installs the malware, which in turn might give an attacker unfettered access to the targeted organization's internal computer network.
- **fraudulent websites and social media**[8]—an exploit that uses a fraudulent website (or social media site such as Facebook) to trick the victim into clicking on a link that downloads malware to the victim's computer. As in baiting, the installed malware may then give an attacker access to the victim's personal information or exploit the victim's computer for fraudulent purposes.
- **pretexting/reverse social engineering**—an exploit that creates and uses a real or an invented scenario (the pretext) to engage a targeted victim in a manner that increases the chance the victim will divulge information or perform other actions that would be unlikely in ordinary circumstances. A sophisticated example of pretexting is reverse social engineering, which was described above in the context of nontechnical social engineering scams. When applied to technical (online) interactions, reverse social engineering has proven to be a very effective computer-based exploit.
- **phishing/spear phishing**—an exploit generally defined as a phisher impersonating a trusted third party to gain access to private data. Typically, the phisher sends an email that appears to come from a legitimate business or individual (e.g., a bank, credit card company, or fellow employee) requesting verification of information and warning of dire consequence if it is not provided. The email usually contains a link to a fraudulent webpage that appears legitimate—sometimes with company logos and content—and requests that the victim provide private information (e.g., Social Security number, bank account number, or banking PIN). Social engineering, and particularly phishing, has become more sophisticated over time: Attackers

---

[7] Reverse social engineering is also prevalent in social engineering exploits that use technical means, as we discuss later.

[8] Fraudulent websites are also referred to as *phishing websites* [APWG 2005].

learn which techniques are most effective and alter their strategies accordingly [Downs 2006, 2007]. An example is *spear phishing*, a form of phishing in which the attacker initially gathers personal information about the target victim and uses it to tailor the phishing scheme, which increases the probability of success [O'Brien 2005].

Table 1 summarizes the salient characteristics of social engineering attacks, typical information sought, and possible consequences of the incident. The information sought and potential outcomes are, not surprisingly, much the same as the targeted information and consequences in cyberattacks generally, although the methods of attack differ somewhat, especially regarding salient characteristics in the first column of the table. These characteristics generally inform our approach to describing social engineering incidents and identifying patterns in these attacks.

Table 1: Summary of Social Engineering Characteristics

Salient Characteristics	Typical Information Requested	Potential Consequences/Outcome
Appeal • usually good news or bad news • sense of urgency • sensitive or confidential matter • impersonating known sender Desired response • provide specific information • update personal/account information • click on link in email message • open an attachment Suspicious indicators • generic greetings • suspicious context • poor grammar or spelling • strange or unusual sender • incorrect information • illegitimate embedded URLs	• account information • user name • password and PIN • credit card number • Social Security number • bank account number • bank routing number • email address • telephone number • other personal information	• financial loss • identity theft • personal, confidential, or proprietary information stolen • intellectual property stolen • computer compromised, malware or virus implanted • data, software, and/or hardware assets manipulated or destroyed • personal or organizational embarrassment • political gain • denial of service

4 Review of Research on Social Engineering UIT Incidents

In the current phase of our work, we focused our efforts on social engineering UIT incidents, building upon research reviewed in the Phase 1 UIT project. This analysis, along with examination of case studies (discussed in Section 5), helped to refine the broad list of contributing factors developed in the initial phase of work.

For the current phase of our work, we reorganized the list of possible contributing factors from Phase 1 into three broad categories: demographic, organizational, and human. Exemplar factors in each of these three categories are discussed below.

Demographic Factors[9] (relating to UIT victim or organization)

- gender—Research has sought to determine if social engineering susceptibility differs between males and females. Maintaining a database that tracks these data ultimately will allow analysis of possible gender differences.
- age—Research has sought to determine if age-related differences exist in social engineering susceptibility.
- personality traits[10]—A limited amount of research has attempted to identify possible correlations between certain personality traits and social engineering susceptibility.
- cultural factors[11]—There has been little research aimed at identifying possible cultural differences in social engineering susceptibility. We briefly discuss possible cultural influences.

Organizational Factors[12]

- inadequate management or management systems—Inadequate management or management practices can increase organizational vulnerabilities to social engineering exploits. Management must enable a culture of network safety and security through management practices including
  - adequate staffing of individuals who work collectively to protect the network and reduce network vulnerability
  - adoption of security-related training practices
  - adequate resources to effectively complete task work

---

9 These are within-person factors that characterize who people are and their past experiences.

10 *Personality traits* are stable, inherent aspects or characteristics of a person's personality (e.g., neuroticism, agreeableness).

11 Cultural factors include characteristics of the individual's attitudes and ways of experiencing life that the individual adopted.

12 Compare this categorization of organizational factors with the categories identified in the report on Phase 1 [CERT 2013]: data flow, work setting, work planning and control, and employee readiness. As noted, we determined that this classification may not be readily applied to UIT incidents, so the current list of organizational factors uses different categories. However, each of the workplace factors in the original set still appears in our list, albeit embedded within definitions or descriptions of other factors. We believe that the revised set will be more appropriate for guiding the coding of incident data into the UIT database.

  - clear communication and dissemination of timely security information from the top down and bottom up such that information about potential exploits and vulnerabilities is circulated
  - implementation of mitigating strategies and actions that are not only encouraged but well planned and rehearsed
- insufficient security systems, policies, or practices—Insufficient or inadequate security systems, policies, or practices (i.e., relaxed stance on security and neglect of security norms) may foster workforce complacency with respect to security practices.
- job pressure—Organization-imposed performance pressures (i.e., difficult performance objectives or expectations, time constraints, unrealistic task difficulty, and high task load) may adversely impact human performance.

Human Factors[13]

- lack of attention—An individual is preoccupied or does not pay sufficient attention to the task (distraction). An individual may also have change blindness (i.e., the inability to detect changing cues) or simply lack awareness of the situation and contextual or physical cues associated with suspicious activities.
- lack of knowledge/memory failure—An individual is ill-prepared to recognize cues because of knowledge gaps (e.g., failure to recognize features of fraudulent situations) or memory failure (e.g., inability to recall appropriate security procedures or recalling the incorrect security procedures).
- faulty reasoning/judgment—An individual exhibits incorrect reasoning or judgment or may devote insufficient cognitive resources for correct reasoning and judgment.
- risk tolerance/poor risk perception—A high-risk-taking or risk-tolerant individual may exhibit risky behavior despite cybersecurity training, while a risk-averse individual may be less likely to knowingly take risky actions. In addition, an individual may habituate to repeated system warnings.
- casual values/attitudes about compliance—An individual may have a casual attitude about the importance of complying with security policies and procedures. Employee attitudes, normative beliefs, or habits may have influenced an individual's lack of compliance with information-system security policies.
- stress/anxiety—Subjective mental stress because of workplace conditions, such as heavy or prolonged workload and constant time pressure, may be correlated with higher task error rates.
- physical impairments—Physical states (i.e., fatigue, illness, injury, and side effects of drugs) may adversely impact human performance as well as other cognitive states such as attention, memory, and reasoning.

The tables in Appendix A summarize the research findings, which are discussed in some detail in this section. The refined set of factors also serves to inform additions and modifications made to

[13] In the initial study, employee readiness factors (e.g., inattention, stress, drug side effects) were listed under organizational factors. In the approach described above, we determined that these factors fit more naturally in the human factors category.

the CERT insider threat database, which is being used to support our incident data collection efforts. Corresponding changes have been made to the database to allow for more useful and accurate representation of UIT cases. We intend to use this database to identify trends and patterns across social engineering and UIT cases within individual time slices and across time. This information could increase our understanding of potential UIT causes and facilitate identification of effective mitigation strategies.

4.1 Research on Demographic Factors

Several studies examined possible associations of various demographic factors with social engineering susceptibility. Unfortunately, findings varied on most demographic variables, as did the quality of the scientific research employed in the studies. Research on demographic factors must be characterized as largely inconclusive at this time.

4.1.1 Gender

A survey conducted by Carnegie Mellon University researchers [Sheng 2010] studied the relationship between demographics and phishing susceptibility. A role-playing survey was administered to 1,001 online survey respondents. Results indicated that females were more susceptible than males to phishing. Similar gender differences were reported by Halevi and colleagues, who suggested that women may feel more comfortable with digital communication and may be more inclined to reply to emails that advertise commercial offers or prizes [Halevi 2013]. In contrast, a large-scale phishing experiment conducted with more than 10,000 human subjects in a university setting found no significant gender-related patterns in phishing susceptibility [Mohebzada 2012]. In an initial phishing attack involving spoofed email that navigates the user to a website to change a student's password, males and females were equally deceived; in a second phase of the attack that used a survey to harvest personal information, nearly 61% of the victims were male compared to only 39% females. These studies used different methodological approaches, ranging from survey studies to empirical studies, and there are differences in the extent to which confounding variables like experience and course of study or job position may have been controlled.

4.1.2 Age

In general, several studies found a significant negative correlation between age and phishing susceptibility (e.g., Sheng 2010, Jagatic 2007). The role-playing survey study conducted by Sheng and colleagues found that participants between the ages of 18 and 25 were more susceptible to phishing than other age groups (26–35, 36–45, 46–55, and older than 56) [Sheng 2010]. An overall success rate of 72% resulted from a phishing experiment with 487 students at Indiana University [Jagatic 2007]; ages ranged from 18 to 24 years, with a slightly higher susceptibility in younger students. (Note that this age range falls entirely within the youngest age category defined by Sheng and colleagues [Sheng 2010].) On the other hand, Mohebzada's phishing experiment found no evidence for age-related patterns in phishing susceptibility for students at different undergraduate levels (freshman, sophomore, junior, senior) [Mohebzada 2012]; again, the age range in this study is more restrictive than that in Sheng and colleagues' study [Sheng 2010]. Furthermore, Dhamija and colleagues found no significant differences in phishing susceptibility between students, faculty, and staff in a university setting [Dhamija 2006].

In describing the contradictions between their results and the results of others (such as Sheng and colleagues), Mohebzada and colleagues observed that one possible reason was that their study, which mimicked real-life phishing attacks, was more realistic than surveys [Mohebzada 2012]. This may be true, but as pointed out above, the lack of a significant trend within the youngest category (18–24) does not diminish the possible existence of an age-related trend spanning the broader range between 18 and 56+ years. The lack of significant results obtained by Dhamija and colleagues [Dhamija 2006] does cast some additional doubt on the existence of an age-related association. It is possible that other factors may be at the root of the relationship, such as amount of experience (which is discussed in Section 4.3, Research on Human Factors).

4.1.3 Personality Traits

Some researchers believe that personality traits may play a role in susceptibility to social engineering exploits [Alseadoon 2012, Parrish 2009, Halevi 2013]. Differences in personality may influence the manner in which people interact with others, approach decisions, respond to job uncertainties or job pressures, and react to social engineering exploits.

Contemporary personality theory classifies humans on five broad personality dimensions or traits (also called the Big Five personality factors). The common Big Five factor model [Digman 1990] includes neuroticism, extraversion, openness to experience, agreeableness, and conscientiousness (also used, for example, in the work of McCrae and John [McCrae 1992] and Weiner and Greene [Weiner 2008]), which are defined below.

- *Neuroticism* is the tendency to experience unpleasant emotions easily, such as anger, anxiety, depression, or vulnerability. It is sometimes called emotional instability, and people who score high on neuroticism are emotionally reactive and vulnerable to stress (lacking the ability to cope effectively with stress, they may have a diminished ability to think clearly and make decisions). In contrast, people who score low on neuroticism tend to be more calm, emotionally stable, and free from persistent negative feelings. A study of phishing susceptibility and the Big Five personality traits found that neuroticism was most highly correlated to responding to a phishing email scheme [Halevi 2013].
- *Extraversion* is the tendency to seek out the company of others; extroverts enjoy interacting with people and are perceived as being enthusiastic, action oriented, and full of energy. Extraverted personalities often seek excitement and tend to be assertive. Introverts have lower social engagement and energy levels than extraverts: They tend to seem quiet, low-key, deliberate, and less involved in the social world. Introverts are not necessarily shy or antisocial; rather they are more independent of their social world than extraverts. Parrish and colleagues [Parrish 2009] suggest that extraversion can lead to increased phishing vulnerability, and they cite empirical research that found that high extraversion was associated with people giving up sensitive information (to gain acceptance to a social group).
- *Openness* is associated with intellectual curiosity, creativity, an appreciation for different ideas and beliefs, a willingness to try new things, and the desire to seek out new experiences without anxiety. People with low scores on openness tend to have more conventional, traditional interests, and they tend to be conservative and resistant to change. Parrish and colleagues speculated that because openness is associated with technological experience and computer proficiency, people who score high on openness could be less susceptible to social

engineering attacks; on the other hand, they suggested that a general openness to all experiences and tendency toward fantasy could play into the criminal's hands [Parrish 2009]. Two empirical studies tend to favor the hypothesis that openness contributes to social engineering susceptibility. A study with 200 Saudi Arabian students found a significant relationship between individuals scoring high on the openness personality trait and responding to a phishing email attack [Alseadoon 2012]. Another study found that people who scored high on the openness personality factor post more information on Facebook and use less strict privacy settings [Halevi 2013].

- *Agreeableness* is a tendency to be compassionate and cooperative rather than suspicious and antagonistic toward others. The trait reflects individual differences in general concern for social harmony. Agreeable individuals value getting along with others and are generally considerate, friendly, generous, helpful, and willing to compromise their interests with others. Agreeable people also have an optimistic view of human nature. Agreeableness is positively correlated with good teamwork skills, but it is negatively correlated with leadership skills. In contrast, a person who scores low on agreeableness may place self-interest above getting along with others. Less agreeable people tend to be distant, unfriendly, and uncooperative, and their skepticism about others' motives might cause them to be suspicious. This trait may be the one most highly associated with social engineering susceptibility: Facets of agreeableness that would seem to be most vulnerable to phishing exploits are trust, altruism, and compliance [Parrish 2009].
- *Conscientiousness* focuses on self-discipline, dutiful action, and a respect for standards and procedures. This trait shows a preference for planned rather than spontaneous behavior. People who score high on conscientiousness tend to be known for their prudence and common sense. People who score low on conscientiousness are typically more impulsive and spontaneous. People who are high in conscientiousness tend to take longer to make a decision; those low in conscientiousness are more likely to make a snap decision. Presumably, higher levels of conscientiousness would make individuals more likely to follow training guidelines and less likely to break security policies [Parrish 2009]. Consistent with this view, a study demonstrated that low levels of conscientiousness predicted deviant workplace behavior such as breaking rules or behaving irresponsibly [Salgado 2002].

There is some disagreement among researchers about the names and definitions of the five personality traits as well as what personality inventory tests appropriately measure them [Goldberg 1971, John 1999]. Regardless, some evidence suggests that no matter what these traits are, they are somewhat intercorrelated [Goldberg 2006].

A caveat is that these five personality traits may not reliably predict behavior; rather the facets that compose these traits may be more predictive than the traits themselves [Paunonen 2001]. For example, a study by Workman investigated the relationship between phishing susceptibility and six personality constructs (based on work by Cialdini [Cialdini 2001]) [Workman 2008]. As shown in Table 4, five of the six factors studied yielded significant correlations with susceptibility to social engineering attacks [Workman 2008]. For example, people who are higher in normative commitment (tendency to form implied obligations to others), more trusting, and more obedient to

authority are more likely to succumb to social engineering attacks. The reactance factor was also positively correlated with susceptibility but not at a statistically significant level.[14]

Table 2: Social Engineering Factors Studied by Workman

Factors	Constructs	Examples
Normative commitment*	C1: Reciprocation as obligation	Free samples
Continuance commitment*	C2: Cognitive investment and perceptual consistency	Spending money on losing ventures
Affective commitment*	C3: Social "proof" as behavioral modeling and conformance	Imitating celebrities
Trust*	C4: Likeability and credibility	Trusting sports figures
Fear*	C5: Obedience to authority and acquiescence to threat of punishment or negative consequences	Obeying commands to avoid humiliation
Reactance	C6: Scarcity and impulsivity	Placing greater value on perceived scarce items

*Factor correlated significantly with susceptibility to social engineering attacks [Workman 2008].

4.1.4 Culture

We found little published research that addresses possible cultural differences in susceptibility to social engineering exploits. None of the social engineering UIT research that we found specifically employed comparative studies across defined cultural variables. At best, one can only draw tentative conclusions in comparing the few experiments conducted in non-Western cultures with those reported from Western countries. Those results suggest that there is little, if any, difference in phishing susceptibility, at least between the Western and Middle Eastern populations used in these studies. The experiment performed by Mohebzada and colleagues [Mohebzada 2012] took place in the Middle East, with participants sampled from the American University of Sarjah in the United Arab Emirates. Mohebzada and colleagues reported that 8.74% of the sample of 10,917 students, faculty, and staff fell for the initial phishing exploit. Students were found to be more susceptible to phishing attacks than faculty or staff, and warning notices against phishing attempts were largely ignored. Consistent with other findings reported in current literature [Dhamija 2006; Downs 2006, 2007; Sheng 2007], users had difficulty recognizing the phishing schemes. Similarly, a study conducted in Saudi Arabia with 200 students reported a 7% response rate to the phishing email [Alseadoon 2012]. These statistics on phishing susceptibility are in line with published results from a variety of studies reporting response ranges between 3% and 11% in Western cultures [Dhamija 2006, Jakobsson 2006, Knight 2004, Mohebzada 2012].

4.1.5 Summary

In summary, there is limited support for the notion of individual differences in phishing susceptibility across demographic factors of age, gender, or personality, and possible cultural differences have not been sufficiently studied. Because of possible methodological problems in certain studies of age effects, the role of age in social engineering susceptibility may be confounded with related factors such as amount of experience. Further research is needed to resolve some of the methodological uncertainties and apparent contradictions in findings among published studies to date. While there is limited research on personality traits, there is some

[14] The personality constructs examined by Workman and other derivative works served to inform the development of the social engineering tactics taxonomy described later in this report (see Section 6.5).

evidence to suggest that it may be possible to use personality profiles to identify individuals who are at a higher risk of falling for social engineering scams. There have been no formal studies of possible cultural differences, but research to date does not indicate any cultural differences in response rates to social engineering scams. Despite the lack of strong relationships between social engineering susceptibility and various demographic factors, we believe that it will be useful to continue to record demographic information as case studies are tabulated and entered into the UIT database. Besides tracking personal characteristics such as those described here, records should be kept relating to the individual's role in the organization (e.g., position title) and the organization's industry sector.

4.2 Research on Organizational Factors

Organizational factors refer to management practices, policies, work environment, workload, and related aspects of the workplace that may contribute to performance deficiencies and human error, which in turn underlie certain types of UIT incidents. Direct mention of such organizational factors in published research within the cybersecurity and insider threat domain is rare, although these factors play a prominent role in the scientific literature on safety and human error (e.g., Dekker 2002).

One possible reason why the deep-seated organizational factors tend to be overlooked in investigations of incidents (e.g., accidents, security incidents, insider threat incidents) is that the immediate cause or point of failure (such as a human error) is the easiest to identify; the organizational factors are more abstract and harder to identify because their failures are often nested in broader organizational aspects such as team management, company policies, company enforcement of policies, and management systems and practices. Poor workplace conditions that produce human errors or deficiencies in human performance may be described in terms of a variety of issues, including poor communications relating to a task and its goals, confusing procedures or directions, faulty design of systems that reduce usability (e.g., lack of appropriate feedback), inadequate resources to accomplish a task, environmental stressors (noise, temperature), changes in routine or job pressures due to unrealistic task deadlines, or poor security practices or systems (e.g., as shown in the work of Pond and Leifheit [Pond 2003]). Underlying causes of these conditions often may be traced to management practices or management systems, so we have classified these factors as organizational issues.

While research (e.g., the work of Pond and Leifheit [Pond 2003]) aimed at defining causal factors implicated in human error investigations typically cites a large number of possible contributing factors, we have found that information available on UIT incidents generally is not sufficiently detailed to enable such fine distinctions. Therefore, we reorganized and recategorized our original list of organizational factors into general types that appear to be most useful in guiding our efforts in data collection and tracking of UIT incidents.

4.2.1 Inadequate Management and Management Systems

We have defined the broad category of inadequate management and management systems to encompass many organizational pitfalls that increase the likelihood of an individual making an error. Effective management includes practices to ensure the availability of qualified staff, assignment of tasks to staff who have appropriate capabilities and experience, and availability of materials and resources to complete the task [Leka 2004]. Failures in any of these areas can

produce job conditions that breed employee dissatisfaction, stress, and attitude problems (such as disgruntlement). The following are examples, adapted from the work of Pond and Leifheit [Pond 2003], of management and management systems that not only reduce productivity and job satisfaction but also create conditions that promote human error:

- poor communication related to the task
- confusing procedures or directions
- tools or systems with design deficiencies (such as poor user interfaces and inadequate system feedback or status)
- problems with the work environment (e.g., noisy, hot, cold)
- inadequate materials or resources (insufficient resources to successfully and efficiently complete the job)

Most of these conditions have multiple deleterious effects on employee job performance and morale; a particularly harmful effect is increasing job stress [Leka 2004]. In turn, stress negatively impacts cognitive processes (described later in Section 4.3, Human Factors).

According to the World Health Organization, research findings show that "the most stressful type of work is that which values excessive demands and pressures that are not matched to workers' knowledge and abilities, where there is little opportunity to exercise any choice or control, and where there is little support from others" [Leka 2004, p. 5]. A general observation from many survey studies of phishing is that people are not highly informed about the nature of the phishing threat or how to recognize social engineering schemes [Dhamija 2006, Mohebzada 2012]. Users who lack knowledge about social engineering schemes such as phishing are more susceptible. For example, Downs and colleagues found that users who could correctly define phishing were less vulnerable to a phishing attack in a role-playing scenario, and participants who had experience with phishing websites were less likely to click on phishing links [Downs 2007]. However, general knowledge about computer risks and concepts (e.g., cookies, spyware, viruses) was not related to phishing susceptibility. Organizations should take care to provide adequate training to raise awareness of social engineering risks.

4.2.2 Insufficient Security Systems, Policies, and Practices

Another consideration relevant to organizational factors is the effectiveness of security practices, policies, and tools. Security practices are often difficult and confusing for an average computer user, and usage errors caused by these difficult security systems can yield serious consequences [Whitten 1999]. In addition, an organization may provide inadequate or ineffective security through its policies (e.g., whether users are required to change passwords periodically) or its technical and defensive measures (such as firewalls or other computer security systems). At the other extreme, security systems, policies, or practices may be too strict or too difficult for most workers to follow, which also may undermine organizational security. Systems that are difficult to understand or use are negatively perceived by users and are less likely to be used [Venkatesh 2003]. Difficulty using security systems may also encourage users to employ shortcuts around these system processes, which may make them more susceptible to UIT incidents. Considerable research indicates that usability and security do not often coexist in computer systems. For example, easy-to-use passwords are not always secure, but secure passwords are often not easy to

use [Zurko 1996]. However, we recognize that system security policies and practices may not protect against the most advanced, highly obfuscated socially engineered messages to users.

4.2.3 Job Pressure

Numerous workplace and environmental conditions have been implicated as sources of employee stress and fatigue. Although the definition of stress varies across research domains, there is broad agreement about conditions that cause stress. Studies consistently reported that time pressure and workload are major sources of stress. Time pressure negatively affects performance of even well-trained individuals [Lehner 1997]. Heavy and prolonged workload can cause worker fatigue, which adversely affects performance [Soetens 1992]. It is clear that stressors in the workplace can adversely impact human performance and error rates.

As we noted in the first phase of this work [CERT 2013], the new view of human error is that when user errors result from deficiencies in system design, it is not sufficient to merely blame the user [Dekker 2002]. As Zurko put it: "When a security breach is said to be caused by 'user error,' the desired implication is that the breach was not the responsibility of the (computer) system, but of the user. Acceptance of that implication is one of the roadblocks to the grand challenge of usable security" [Zurko 2005, p. 189].

If a breach or UIT incident may be traced back to computer system design and usability, inadequate security controls, management practices that increase job pressure, and the like, the problem is an *organizational* one.

4.2.4 Summary

In summary, organizational factors can produce system vulnerabilities that adversaries may exploit, either directly or more typically indirectly, by capitalizing on increased likelihood of human errors and lapses in judgment that place stressed workers at risk of being deceived by social engineering scams. Management and management systems may fail to assign sufficiently qualified personnel to tasks; provide insufficient materials or resources; create inadequate or unusable information-security systems or policies; and present work environments or work planning and control systems that negatively impact employee satisfaction or cause stress that leads to human errors or lapses in judgment.

However, we recognize that organizational factors are difficult to identify as contributing factors to socially engineered exploits and are also difficult to change. For example, the UIT and social engineering class of exploits is evolving so rapidly that organizational policies and practices cannot be created quickly enough to protect organizations. In addition, organizational staffing involves a variety of educational backgrounds, often not from the computer sciences, which can encumber the identification of, and warning communications about, potential exploits. Also, organizations often must balance operational goals (e.g., short product development cycles, multiple product release dates per quarter) with security goals (e.g., protecting intellectual property and other assets from adversaries) to maintain a competitive edge in the market; historically, many organizations have valued operational goals above security goals. These organizational factors that contribute to a rise in social engineering UITs are challenging to address.

4.3 Research on Human Factors

Despite the best organizational efforts to educate users or impose security practices and security control systems and safeguards, social engineering scams, especially phishing schemes, continue to succeed. A number of studies and research papers emphasize the need to better understand the psychological aspects of social engineering exploits—why people fall for these scams—to develop more effective security practices, systems, and training approaches to combat them. Much of the research is focused on phishing and spear phishing exploits, although the findings may be generalizable to social engineering threats. Research suggests that human factors may contribute to increasing human errors in the context of UIT incidents.

4.3.1 Lack of Attention

Dhamija and colleagues studied features of phishing websites to determine what users attended to in assessing the websites' legitimacy [Dhamija 2006]. Participants were shown 20 websites and asked to identify which ones were fraudulent and which were authentic. They found that 23% of the 22 participants ignored browser-based security cues (address bar, status bar, Secure Sockets Layer [SSL] padlock icon); these individuals made incorrect choices 40% of the time. In addition to the problem of lack of attention to security cues, Dhamija and colleagues also found that visual deception practiced by phishers could fool even the most sophisticated users. More recently, Erkkila reported that users may not notice or read security warnings or other security indicators, so they fail to notice the absence of security indicators (e.g., the SSL padlock icon in the status bar) when they should be present [Erkkila 2011].

Studies conducted and reported by Jakobsson examined cues and design issues leading to security decisions in evaluating the legitimacy of email and websites [Jakobsson 2007]. Jakobsson found that spelling was the primary characteristic used to assess email messages' legitimacy. In contrast to commonly held beliefs, Jakobsson and colleagues found that participants do examine URLs (those displayed as mouseovers in email and those displayed in the webpage address bar). Participants are suspicious of spelling errors in these addresses, and they are more suspicious of long URLs than short ones.

Although these studies indicate the possible role of errors in attention or perception in assessing the legitimacy of email and websites, phishing campaigns that are highly obfuscated may trick even the most highly trained individuals into believing a message to be genuine; such campaigns may have no relevant cues upon which to base a judgment that the message or website is illegitimate. In many cases, other, perhaps more subtle and sophisticated factors besides attention and perceptual judgment are at play, such as a sense of urgency. Urgency, which plays on the empathy or inclination of the victim to help someone in need, is a particularly effective characteristic in successful phishing emails [Milletary 2005, Chandrasekaran 2006]. A study by Vishwanath and colleagues suggests that individuals focus disproportionately on urgency cues, often ignoring other elements of the email such as its source, grammar, and spelling [Vishwanath 2011]. Because these other elements aid the detection of deceptive stimuli [Jakobsson 2007], an individual's lack of attention to these elements may increase the individual's susceptibility to phishing. In addition, Vishwanath found that individuals were far more likely to respond to phishing emails when they were faced with large email loads. These results are consistent with research on attention and cognitive load, described in the previous section, showing that high workload narrows attention [Houston 1969, Stokes 1994].

4.3.2 Lack of Knowledge and Memory Failure

Consistent with research that contends users do not notice cues that should reveal suspicious or fraudulent phishing sites, Sharek and colleagues [Sharek 2008] reported that users lack knowledge about design inconsistencies that distinguish real and fake error messages. Users are reported to lack knowledge and basic understanding of the structure of the internet and computer systems in general [Erkkila 2011], although as previously noted, general knowledge about computer risks and concepts (e.g., cookies, spyware, viruses) does not appear to be related to phishing susceptibility [Downs 2007]. Based on research relating attentional processes to phishing susceptibility, key knowledge elements include knowledge about security features and understanding of URL and domain name syntax [Dhamija 2006, Downs 2006]. Also, as noted above, individuals who experience workload or other types of stress are more likely to suffer from attentional or memory deficits that increase vulnerability to social engineering exploits. Research supports the claim that experience *does* have a positive effect: previous exposure to phishing attacks makes users less likely to respond to phishing exploits in the future [Downs 2007].

4.3.3 Faulty Reasoning or Judgment

Errors in judgment and reasoning can occur when the individual experiences cognitive bias. Several types of cognitive bias exist, but the prominent types include attentional bias, memory bias, and decision-making biases. Kahneman and Tversky have shown that people's decisions are often biased and are not purely rational (i.e., all decision options being systematically considered and decisions being made based on factual reasoning) [Kahneman 1979]. An example of decision-making bias occurs when individuals tend to think that threats are highly unlikely (e.g., they underestimate the abilities of social engineering attackers and overestimate the defensive capabilities of organizational security systems) and consequently ignore such threats [Sandouka 2009]. Also, some users feel that use of strong security features will impede their job [Erkkila 2011]. Annoyance with popup messages may actually lead (impatient) users to click on fake popups [Sharek 2008], which contributes to poor judgment in assessing risks.

Research on attentional bias by Jakobsson and colleagues examined some of the thought processes involved in assessing possible threats [Jakobsson 2007]. Importantly, Jakobsson and colleagues reported that people judge *relevance* before *authenticity*; in other words, participants' decision about the legitimacy of an email or website was often based on the content, instead of cues and signs of authenticity. For example, participants considered a website that offered a monetary reward to be "phishy," regardless of whether it appeared authentic. Likewise, participants considered emails that requested passwords up front to be "phishy," whereas they considered emails that only appeared to contain information to be safe. The problem that Jakobsson and colleagues pointed out is that users could be drawn to a site by an email that appears to be for information only, and once at the presumed trustworthy site, they could be asked for credentials or private information.

Watters approaches the problem of phishing by considering the cognition involved in the establishment and assessment of trust and trustworthiness in email messages [Watters 2009]. He argues that over time, positive experiences with email (i.e., absence of negative phishing consequences) tend to gradually build up trust through a habituation process (as described by Staddon and Higa [Staddon 1996]) that desensitizes the user to certain phishing-relevant characteristics of email messages. Because users rely on habit and experience (the vast majority of

which involve scam-free emails and websites), this habituation process tends to decrease the level of cognitive processing of phishing cues so that important, and sometimes obvious, cues are ignored [Watters 2009]. For example, a shallow level of processing might occur when users take a cursory glance at the "Sender" or "Subject" fields of an email and quickly respond by (inappropriately) clicking on the link in the body. Deeper cognitive processing of the message would involve the user "reading the contents carefully, cross-checking the claims made in the e-mail carefully, and then verifying whether the displayed link actually matched the known good link of the service in question" [Watters 2009, p. 4]. Finally, as noted earlier, errors in judgment may also occur as a result of attention or memory impairments brought upon by factors such as workload stress.

4.3.4 Risk Tolerance and Poor Risk Perception

Psychological research on individual differences (i.e., how individuals differ in their behavior) in risk-taking behavior should also be considered when studying UIT. The National Institute of Standards and Technology (NIST) defines risk as the net negative impact of the exercise of vulnerability, considering both the probability and the impact of occurrence [NIST 2002]. From a cognitive process point of view, risk-taking behavior is a function of risk perception (a decision maker's assessment of the risk inherent in a situation), risk propensity (the general tendency either to take or to avoid risk), and a decision process (determining how to act in the face of a risky situation). Considering risk propensity, high-risk or risk-tolerant individuals may take big risks despite cybersecurity training, while risk-averse individuals are less likely to knowingly take risky actions. As we have noted in describing studies of phishing, users largely ignore warning notices against phishing attempts [Mohebzada 2012]. There are a variety of possible reasons, including lack of attention or habituation as described above, lack of patience, and high risk-taking propensity. People who are less risk averse are more likely to fall for phishing schemes; those who are more risk averse are the less likely to do so [Sheng 2010]. Risk perception and risky decision making are important components of a future descriptive cognitive model that may illuminate how the adversary's knowledge of human behavior enables a UIT exploit. Organizations might use this type of model to identify possible mitigation strategies and countermeasures.

4.3.5 Casual Values and Attitudes About Compliance

Employees whose attitudes and beliefs do not align with company security practices and policies and so fail to comply with them are a major threat to information-system security. Employee attitudes (e.g., manner, disposition, feeling, and position, with respect to a person or thing) and normative beliefs (i.e., the perception of what other people believe) can impact the intention to comply with information system security policy [Pahnila 2007, Bulgurcu 2010]. Ultimately, compliance with effective security policy may reduce the incidence of socially engineered UIT exploits.

In one study, sanctions did not significantly influence employees' intention to comply, and awards did not have a significant effect on actual compliance [Pahnila 2007]. A more recent study concluded that attitude toward compliance can be traced back to pre-existing beliefs. For example, beliefs about overall assessment of consequences are immediate antecedents of attitude, so factors that motivate employees to comply with the information-system security policies extend beyond sanctions and rewards [Bulgurcu 2010]. The latter study empirically found that the impact of the

cost of compliance on actual compliance is as strong as the impacts of the benefit of compliance and the cost of noncompliance. This result highlights the importance of the concepts of costs and benefits of compliance and noncompliance in research on attitudes about information security. While rewards do not necessarily lead employees to believe that information-system security policies are mandatory, they influence perceptions of the benefit of compliance, which, in turn, affect employees' attitude toward compliance. Further, because employees perceive information security compliance to be costly (i.e., impeding job-related functions), it is important for organizations to allocate a certain amount of employees' time to fulfilling compliance requirements such that compliance efforts do not compete with daily job functions [Bulgurcu 2010]. Bulgurcu and colleagues believe that creating a security-aware culture within the organization will improve information security. Organizations should provide training and awareness programs to their employees to ensure that they know what they need to do to comply with information security rules and regulations.

4.3.6 Stress and Anxiety

As noted earlier in the discussion of organizational factors, workplace conditions such as heavy or prolonged workload and constant time pressure are sources of stress. However, job-imposed stressors do not necessarily translate to higher levels of internalized stress, or subjective mental workload. Job-imposed time pressures have been found to negatively affect performance of even well-trained individuals [Lehner 1997]. In addition, heavy and prolonged subjective mental workload can cause employee fatigue, which adversely affects performance [Soetens 1992]. Workplace stressors (e.g., organization-imposed time pressures) contributing to higher levels of subjective mental workload tend to negatively impact human performance by, for example, narrowing visual attention such that important cues attributed to malicious activity may be missed [Houston 1969, Stokes 1994] and by reducing cognitive resources needed for effective job performance [Davies 1982, Hockey 1986, Wachtel 1968]. An obvious implication is that reducing work-related stress levels by adjusting time pressure and workload is one way to reduce the likelihood of UIT incidents.

4.3.7 Physical Impairment

While there is no research that specifically relates physical states to social engineering vulnerability, there is substantial evidence that physical states[15] may impact human performance. Substance abuse may negatively affect cognitive functioning. For example, a study of neurocognitive impairment reported impaired neurocognitive performance in approximately two-thirds of patients who entered a 14-day inpatient substance abuse unit; the most frequently compromised areas of functioning involved attention, memory, calculation, abstraction, ability to follow complex commands, and visuospatial skills [Meek 1989]. Abuse of drugs and alcohol may be associated with loss of productivity, among other problems [HealthyPeople.gov 2013]. Drug effects are not confined to substance abusers; people who suffer from physical injuries or illnesses may take prescription drugs that have deleterious effects on cognitive performance (e.g., judgment, memory, risk-taking behavior). A drug may lower an individual's risk threshold by lowering inhibition or lowering risk perception sensitivity (e.g., might increase aggression and distract someone from perceiving a risk). Hormones, particularly dopamine, can also affect the

[15] By *physical states*, we mean conditions such as fatigue, illness and injury, and the effects of drugs or hormone imbalance.

amount of risks that people take [Park 2008]. Zald and colleagues found that dopamine is more pervasive in the brains of risk-takers, or that they have fewer dopamine-inhibiting receptors [Zald 2008]. They conclude that people with these higher levels of dopamine are more likely to take risks such as abusing drugs and other unsafe behaviors. The implications for drug education and policies are evident (zero tolerance, along with available rehabilitation employee-assistance programs). Such policies should encourage a drug-free environment that reduces the deleterious effects of drug use on risk-taking behavior.

4.3.8 Summary

In summary, many studies emphasize the need to better understand the psychological aspects of social engineering exploits in order to develop more effective security practices, systems, and training approaches to combat social engineering. Some social engineering campaigns may be so well crafted that individuals may still be exploited no matter what countermeasures (e.g., training, policies, etc.) are employed. For the less sophisticated campaigns that offer perceptible cues that a message is potentially exploitive, some of the human factors discussed above may predict the probability of being exploited. Several studies reported that users tend to ignore or do not recognize cues that a particular socially engineered message is malicious. A possible reason includes a lack of attention to these cues or a lack of knowledge about the exploitive nature of the message. In addition, the narrowing of attention can be exacerbated by high cognitive load (high subjective mental workload). There is a need for more research to identify other possible explanations for this result.

Regardless, phishers exploit these cognitive limitations of network users through visual deception to spoof legitimate email messages or websites. In addition, phishing schemes exploit a tendency for humans to focus disproportionately on urgency cues (i.e., the message urges the reader to act quickly). Susceptibility to social engineering attacks also may be traced to problems with poor judgment or cognitive biases: people sometimes underestimate the likelihood of the threats and thus ignore them. Because the vast majority of email and online experiences are scam free, people can habituate to cues and consequently miss the phishing cues, a common phenomenon under conditions of high workload.

Risk tolerance and risk perception represent other significant human factors to be considered in addressing social engineering threats. Research has shown a negative correlation between risk tolerance and susceptibility to phishing, such that people who are less risk averse are more likely to fall for phishing schemes. Another factor to consider relates to the values and attitudes of employees about information security: Because employees often perceive information-security compliance as costly (i.e., interfering with job functions), it is important for organizations to allocate a certain amount of employees' time to fulfilling the compliance requirements. Finally, because work-related stress has deleterious effects on cognitive processes and human performance that may lead to human errors and UIT incidents, organizations should apply effective management practices to create work environments that minimize stress (e.g., minimizing time pressure and optimizing workload).

The use of deception and obfuscation in social engineering UIT incidents, particularly phishing, presents special challenges for research aimed at developing effective mitigation strategies. Deceptive practices that exploit human psychological limitations and vulnerabilities are all the more challenging because the adversaries continue to change tactics. No matter how skilled,

savvy, or trained an organization's employees are, there will always be a chance that a phishing campaign will succeed, especially because it takes only one individual to succumb to the scam to open new opportunities for the social engineer attacker to execute further exploits against the organization. Thus, the research community and responsible organizations and stakeholders are obligated to continue research and information gathering to inform the development of effective training and mitigation tools. Indeed, one implication of the increasing sophistication of social engineering attacks is the need to continue to examine these threats so that new information can be incorporated into updates of training and mitigation strategies. The next section provides a current status update on characteristics and patterns that we have observed to date, based on a small but growing collection of social engineering UIT case studies.

5 Summary of Collected Cases

Case study research helps identify concepts and factors associated with a phenomenon of interest. Though case studies do not constitute a valid research method for making generalizable inferences, without them researchers are left to infer what factors and parameters are important. Collecting and analyzing UIT social engineering case studies are helpful for identifying factors and relationships that may be addressed later in experimental and observational research. Those activities also enable statistical testing of hypothesized relationships (e.g., causal, correlational, moderating, mediating, predictive) between factors and incidents. By informing experimental and observational research, case study research improves the validity and generalizability of these hypothesized relationships.

As we found in the first phase of our work [CERT 2013], social engineering exploits constitute a subcategory of the UIT-HACK vector. To better understand the scope and variety of social engineering exploits, we use a case study approach that collects, abstracts, and reports on actual incidents. Using a set of descriptive parameters borrowed from Phase 1 research, we summarized incidents succinctly and expressed them in a standardized manner for informal review. These parameters are defined using the following incident template:

- INCIDENT ID: <assigned ID number for the case>
- INDUSTRY: <classification of organization>
- STAGING: <single, multiple>
- INCIDENT: <description of the how social engineering was used, across multiple stages where applicable>
- BREACH: <type of loss or compromise>
- OUTCOME: <organizational status resulting from breach>
- RESPONSE: <specific action taken in response to the breach>
- REFERENCES:[16] <URLs or references to sources of incident descriptions>

We use these parameters to summarize representative examples of incidents in the following section and in the social engineering case studies in Appendix B.

5.1 Representative Cases

In this section, we use the above incident template to describe selected UIT social engineering incidents. Appendix B provides a full listing of UIT social engineering cases collected to date. The case information comes from three sources:

- reports captured during the foundational study that fall into the social engineering category
- new reports captured through internet searches using search strings to locate relevant incidents

[16] To preserve privacy and anonymity of the organization, we do not divulge names or identifiable information, including website URLs and citations of news articles or legal judgments. We have omitted the References field from all incident summaries.

- reports referenced in the literature that were subsequently investigated for this study

For ease of presentation and to help reveal certain patterns (described in Section 6), the cases are categorized into single- or multiple-stage attacks. This categorization reflects our observation, based on examining cases collected, that many of the incidents may be decomposed into separate stages that share certain common characteristics that make up patterns or building blocks of incidents.

Information relating to possible contributing factors (behavioral or technical) is generally not made public and is difficult to obtain. In some cases we gathered the information by carefully examining numerous separate information sources (news articles, court records, etc.). The annotations following each example summarize the identifiable contributing factors, which appear in the contributing factors table in Appendix A.

For the purposes of this report, we have included social engineering exploits against clients or users of an organization. Although clients or users may be considered to have a business relationship with the organization, they would not necessarily be considered as organizational insiders. Therefore, an argument may be made to exclude cases that take advantage of an organization's clients (e.g., banking customers). On the other hand, organizations have a vested interest in discouraging or preventing social engineering attacks aimed at their customers—these attacks can damage the organization's reputation and cause loss of customers and revenue. Thus, organizations may take steps to help prevent or combat social engineering threats to information security, such as by informing customers about these threats, how to recognize such threats, and clarification about their privacy and security policies (including identifying the kind of information that is requested from clients via email and information such as passwords, which are never requested in that way). Because the research and concepts discussed in this report are relevant to the problem of social engineering exploits against an organization's clients or customers, we have included cases of this nature in our database.

5.1.1 Single-Stage Phishing Attacks

In the case summarized in Figure 2, the targets of the exploit had all been trained in identifying and resisting phishing attempts after a previous, similar attack. However, the phisher was able to provide a very realistic email (*high obfuscation*) to entice potential UITs, and about five staff members succumbed. The breach involved lists of visitors and their identifying information, so this constituted a serious security threat. However, the organization was able to resist repeated attempts to access more secure types of information.

Incident ID 24

INCIDENT ID: 24

INDUSTRY: Government

STAGING: Single

INCIDENT: Employees were duped by a phishing email about HR benefits that exploited a zero-day vulnerability and downloaded malicious code. The malware masked itself on systems and was designed to erase itself if it tried to compromise a system and was unsuccessful.

BREACH: Only a few megabytes of encrypted data were stolen, but the organization failed to recognize additional dormant, malicious code.

OUTCOME: The organization was forced to disconnect internet access after administrators discovered data being externally siphoned from a server. After initial shutdown, the organization restored external email access but prohibited attachments.

RESPONSE: This was the second widespread social engineering attack. The organization implemented extensive training after the first. The specific response to this incident is unknown.

Figure 2: Single-Stage Phishing Attack, Example 1.

Hijacking of Twitter accounts has become commonplace, and other social media (e.g., Facebook, LinkedIn) are frequent targets of cyberattacks. In addition to offering access to various systems and accounts, they provide background used as intelligence to support the initial phishing. The news organization affected by the attack summarized in Figure 3 had intended to use Twitter for news gathering and to combat rumors, but Twitter's security weakness makes it a prime target. The need for hot-list items and other immediate news information may cause otherwise security-conscious users to relax their guards.

Incident ID 15

INCIDENT ID: 15

INDUSTRY: Information and telecommunication

STAGING: Single

INCIDENT: Attackers sent an innocent-looking email to news service staffers urging them to click on a link to an important article on another news organization's blog that, unknown to the victims, would infect their computers with malware. The malware allowed the hackers to capture passwords to the news service's Twitter account.

BREACH: Access to the news service's Twitter account allowed the attacker to send an erroneous Tweet warning of two explosions in a government building.

OUTCOME: Within minutes, the bogus story had a brief but very real effect on the stock market, causing it to drop significantly. This stock market loss was made up after the story was confirmed to be false.

RESPONSE: This was the second widespread social engineering attack on the news service, which had implemented extensive training after the first. This latest incident happened even though the news service had sent a message warning staffers of bogus emails that were being sent out. After learning the erroneous Tweet caused the stock market to drop, the news organization had all the staffers change their passwords, and it shut down its compromised Twitter account.

Figure 3: Single-Stage Phishing Attack, Example 2.[17]

The case summarized in Figure 4 illustrates that not all social engineering incidents require direct interactions between the UIT and the attacker. In this case, the attackers created a website to entice the developers, hoping they would accept the bait and install the Java plug-in. While the details of the breach were not publicly disclosed, the involvement of law enforcement in the

[17] Figure 15, in Section 6, diagrams this example (p. 43).

incident indicates a serious breach. In addition, the incident involved six UITs within one company and, possibly, developers from other companies. This attack might be an example of reverse social engineering, where the phisher impersonates a provider of technical insight to developers.

Incident ID 743

INCIDENT ID: 1

INDUSTRY: Computer manufacturer

STAGING: Single

INCIDENT: Malware to attack computer manufacturers was spread through a website for software developers. The website advertised a Java plug-in that could be installed on desktops.

BREACH: A few employees of one reported company installed the so-called Java plug-in, which was in fact cleverly placed malware. The incident affected a small number of systems.

OUTCOME: The manufacturer worked with law enforcement to find the source of the malware. The manufacturer's native antimalware software was able to catch the malware and isolate it.

RESPONSE: The affected systems were isolated from the network. The company released a tool to remove the malware.

Figure 4: Single-Stage Phishing Attack, Example 3.

5.1.2 Multiple-Stage Phishing Attacks

As in our research on Example 3 of a single-stage attack, the available information relating to possible contributing factors was difficult to obtain and was gathered by carefully examining numerous, separate information sources. The case summarized in Figure 5 resulted in a lawsuit with considerable numbers of court filings of documents and testimony from both the bank and the manufacturing firm. Details about this attack are available, and a thorough study of this case illuminates the nature of many types of phishing exploits and insider responses.

Incident ID 5

INCIDENT ID: 5

INDUSTRY: Banking and finance, manufacturing

STAGING: Multiple

INCIDENT: The phisher impersonated the company's bank, requesting information to address security concerns. The insider clicked on a link in a phishing email and entered confidential information.

Stage 1 - phishing to multiple bank customers

Stage 2 - spear phishing to executives with likely wire-transfer authority

BREACH: The disclosure included credentials and passwords that enabled outsiders to transfer funds to accounts in several countries.

OUTCOME: The bank was able to reverse 70 percent of total money lost.

RESPONSE: The company recovered the remainder in a court settlement resulting from a lawsuit brought against the bank.

Figure 5: Multiple-Stage Phishing Attack, Example 1.[18]

[18] Figure 16, in Section 6, diagrams this example (p. 44).

5.2 Characterization of Case Study Data

Cases can be compared on factors (i.e., demographic, organizational, and human factors), as documented in Section 4, or on the parameters in the incident template shown in the beginning of Section 5.

In all, there are 28 cases in our UIT social engineering database. All the cases were found online, such as through search engines. Three of the cases (10.7%) have more than one source reference. A breakdown of the sources is as follows:

- news articles: 25/28 (89.3%)
- journal publications: 1/28 (3.6%)
- blog: 1/28 (3.6%)
- other: 1/28 (3.6%)

Media reports are the primary source for the case study data, so our investigation is limited to information provided to those sources or assembled from them. As a result, it is challenging to extract or infer information that reveals possible contributing factors, either behavioral or technical. The following represents our current understanding and characterization of the cases so far.

5.2.1 Demographic, Organizational, and Human Factors

In Section 4 we described conclusions and implications from diverse fields of study and research approaches that examined possible causal factors in human error as well as more directly applicable UIT social engineering incidents. We reviewed all the case study data to determine which, if any, of these contributing factors might be implicated in operational accounts of UIT social engineering incidents. However, we found little relevant information on demographic, organizational, and human factors in published case studies. While the insider threat database allows third-party coders to include information on these types of factors, the coders either inferred this information or did not provide it. This was in part because the references provided for each case study did not include tangible information about these parameters. The insider threat database is currently being revised to capture what tangible information was provided by the references, what information was inferred by third-party coders, and what information was unknown. This information is critical to reducing measurement error in any future statistical analysis that uses database codes.

Out of the 30 case studies of UIT incidents involving social engineering, we found the following information about contributing factors (see also Appendix A):

Demographic

- gender—The gender of victims is stated directly in some of the case study reports. In others, we inferred the gender of the victim based on the case description. Some cases involved both male and female victims. Many of the attacks on financial institutions identified the victim as male, but the data captured does not allow for conclusions about susceptibility based on gender.
- age—Again, we can make inferences based on the type of organization. For software development groups, victims would likely be in their 20s and 30s. The ages of mid-level

financial or government victims are probably somewhat higher. However, no conclusions can be reached based on the case studies we examined.

Organizational

- security systems, policies, and practices—Many of the case studies provide a look into organizational policies and procedures. Some indicate that the victims violated those policies, but most incident summaries do not provide sufficient information to determine whether these factors are involved (organizations generally do not have an automated means of tracking employee actions to audit such actions or to warn employees of possible violations).
- management and management systems—Many of the cases reveal that a simple login identification and password provided the attackers with access to internal emails, company data, and even entire computer networks. In one case, the attacker seemed to have attained computer network access directly from the login and did not need to place malware or execute any other indirect attack in order to cause damage. Organizations must regularly perform extensive security audits to determine how best to improve internal controls; they cannot rely on security established during initial installation of a system.
- job pressure—Certain industry categories, such as news services, place a premium on obtaining and distributing information as quickly as possible. Employees of such organizations may be more prone to outside influence from social engineering due to this pressure.

Human Factors

- attention—At least one case study identified fatigue as a contributing factor: The phishing message was received late at night, and the individual responded without completely analyzing the message. A phisher—in this case a spear phisher—may have information about work hours or other conditions that could affect the likelihood of an attack's success.
- knowledge and memory—Many of the case studies included information about prior staff training. Organizations that provide such instruction indicate that, even with training, a large percentage of employees respond to phishing attacks. Constant refreshers or other means should be applied to maintain trainees' knowledge over time.
- reasoning and judgment—Some cases studies indicated that an employee's safeguards were lowered, perhaps because of the realistic nature of the phishing message and pretext created through reverse social engineering (i.e., offers to assist in preventing or addressing outside attacks, solving bank account problems, or supporting system operations).
- stress and anxiety—In one case (Example 1 of the multiple-stage attack), the victim knew that the organization and its customers were receiving phishing emails. This knowledge may have increased his desire to accept an offer of mitigation that appeared legitimate, thought it was actually another phishing attack.

5.2.2 Discussion and Implications of Sample Data Obtained to Date

The lack of some pertinent information in these case studies does not diminish their importance in the study of possible underlying causes and development of effective mitigation strategies. If anything, the paucity of reporting data underlines the need for mechanisms that would facilitate such reporting. One contribution of the present study is to make stakeholders aware of UIT

research suggesting potential roles of organizational and human factors issues underlying UIT and particularly social engineering incidents. This awareness may encourage government and industry to establish a more robust reporting system that includes possible contributing factors, especially data that relates to demographic, organizational, and human factors.

One goal of this effort was to cross-validate prior research findings on factors (e.g., demographic, organizational, and human) that may impact social engineering exploits and incidents of human error with the case studies collected and added to the insider threat database. However, our case study research found little information that is relevant to possible contributing factors described in academic research. We assume there is a disconnect between individuals researching potential causal and correlational factors of social engineering UIT incidents and individuals responsible for case study reporting. Cybersecurity is a fledgling research field, so it follows that there is a dearth of literature that directly studies the relationship between these factors and social engineering exploits. This paucity elucidates a research gap, though some related literature exists to inform future research efforts.

In addition, cybersecurity research is apparently dominated by computer science engineers who may be unfamiliar or not knowledgeable about behavioral sciences research on factors that may impact human error or the frequencies of social engineering exploits. As a result, cybersecurity engineers conducting research are not always informed of related research findings from other disciplines and may never collect relevant case study data. We intend for our research findings to inform those individuals so they will begin collecting information on potential factors identified in behavioral sciences research. If this is successful, future experimental research could determine the viability of those factors to advance our understanding of social engineering exploits and facilitate the creation of effective mitigation strategies.

6 Conceptual Models for Social Engineering Incidents

In this section, we describe the approach we used to gain a better understanding of the problem of social engineering UITs. We also highlight common features across multiple exploits and identify possible mitigation strategies. Our approach used the following conceptual models or analytical methods:

- attack progression analysis (Section 6.1)
- characterization of attack patterns (Section 6.2)
- system dynamics modeling (Section 6.3)
- ontology of social engineering tactics (Section 6.4)

Section 6.5 discusses the approaches' implications for mitigation strategies.

6.1 Attack Progression Analysis

The concept of an attack progression, popularly known as a kill chain, originated in the military as a way of analyzing attacks in the physical world. The insight was that decomposing attacks into a sequence of phases would make them easier to understand and defend against [Myers 2013]. In 2009, Cloppert adopted this technique for use in cybersecurity [Cloppert 2009]. Cloppert analyzed cyberattacks into the now-classic six phases: Reconnaissance, Weaponization, Delivery, Exploitation, Command-and-Control, and Exfiltration (Figure 6). Hutchins, Cloppert, and Amin subsequently articulated attack progression analysis more formally as a way of developing indicators for advanced persistent threat (APT), which describes an adversary with the ability and the intent to persistently and effectively target a specific entity [Hutchins 2011].

The idea behind attack progression analysis was that knowing the phases of the attack can guide the target's defense planning and enable a weakest-link strategy, in which the adversary's attack can be thwarted at any step of the attack progression. More recent variations on the six-phase model permit more flexibility and account for more complexity in the attacks. For example, Secureworks uses a 12-phase kill chain [Bijou 2013]. In Section 6.2.1, we discuss the customized attack progression that we developed for social engineering attacks.

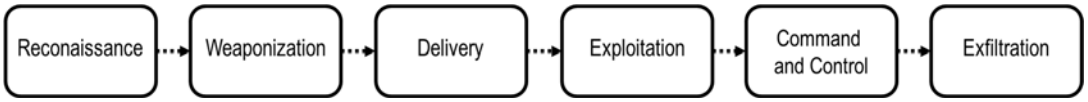


Figure 6: Cloppert's Six-Phase Attack Progression

Other models similar to attack progression have been used for the analysis of social engineering attacks. For example, Laribee and colleagues [Laribee 2006b] present an attack model that includes many of the same aspects as the kill chain, such as research and planning, but provides a more in-depth look at different methods of attack. Other models focus on specific types of phishing attacks, such as internet banking fraud or cross-site scripting attacks [McCombie 2010]. Jakobsson's graph model provides a very flexible way to characterize a wide variety of phishing attacks [Jakobsson 2005].

6.2 Patterns Inferred from UIT Case Studies

Social engineering attacks leverage human psychology and how humans interact with technology. Phishing illustrates how social engineering exploits work. Most phishing attacks have three components: the hook, the lure, and the catch [Parrish 2009]. The hook is the seemingly legitimate email form, website, or mechanism used by the phisher. The lure is the incentive aimed to trick the potential victim into taking the bait. The catch is the information acquired in the attack. Phishing attacks use different approaches toward social engineering of potential unintentional insiders.

Phishing emails can be simple or highly sophisticated. In the simplest cases, the attacker sends out a simple email message that may offer a reward, such as gifts, free trips, or reduced insurance or utility rates. The message generally directs the reader to a URL where the user enters a system password and other login information. In more sophisticated cases, the message may have the look and feel of company letterhead. Again, the company may be a cell phone provider, a bank, or the insider's own organization. The message generally serves the same purpose as the simple email message described above.

Multiple-stage social engineering attacks are common. The first stage uses one of the above methods to obtain account privileges on the UIT's computing resources. The attacker then uses the login information to search the UIT's internal system for detailed information about employees, company policy, or privileged data. The attacker uses insider knowledge about higher level personnel to implement spear phishing attacks. These messages, customized and targeted at individuals rather than large groups, tend to contain information specific to the addressee and to specific internal enterprise conditions. The attacker's goal is to obtain administrator privileges that may allow the attacker to access proprietary data, interfere with internal financial operations, or cause damage to operations through a denial of service or other attacks.

The cases in Section 5 provide examples of these types of social engineering and the resulting chain of events when UITs accepted the bait and returned valuable information to the phisher. One way to characterize these attacks is to describe their general stages and then distinguish classes of social engineering attacks according to patterns evident in the stages. To represent these patterns, we will provide a number of views:

- A *workflow pattern* shows the overall phases of a single- or multiple-stage attack.
- The *use case model* shows these steps as individual use cases and actors.
- The attacker, UITs, messages, and other aspects of the incident are modeled as *classes and subclasses*.
- The *swim lane chart* is a behavioral view that shows use case activities of each actor.
- The *interaction view* shows the collaboration of entities carrying out behaviors in the various swim lanes.

These views not only encompass specific activities, but they also identify the actors, interactions between actors, and the objects exchanged in these interactions. The general patterns provided in this section can be instantiated with specifics and variations for each attack listed in the case studies. In this section, we illustrate the instantiations using Example 2 from Section 5.1.1 (single-stage attack, Case ID #15) and Example 1 from Section 5.1.2 (multiple-stage attack, Case ID #5).

6.2.1 Single-Stage Phishing Attack

The attack progression for a single-stage attack generally comprises five steps, as shown in Figure 7. This is a variation of the kill-chain model discussed in Section 6.1, with some customizations in the delivery, exploitation, and command-and-control steps to accommodate the specifics of social engineering. The steps shown in Figure 7 represent general building blocks on which more complicated attacks may be based. Each phase of the attack has different objectives that can change opportunistically depending on what information is captured during the social engineering operation. The general workflow pattern allows for this flexibility.

Figure 7: Workflow Pattern Showing Phases of a Single-Stage Phishing Attack

In the first phase, the attacker researches possible targets. Based on information gathered, the attacker prepares phishing artifacts. Following this Planning and Preparation phase, the attacker executes the phishing operation by sending phishing emails to recipients in the target organization. While most recipients do not respond, those who do respond become UIT victims. In the Response and Information Capture phase, the UIT unwittingly sends account information to the attacker's system. When this information is received, the attacker conducts the final phase of the attack by using the account access to plant malware or take other measures directed against the UIT or the UIT's organization. Table 3 shows typical actions that characterize each phase of the phishing attack.

Table 3: Steps in a Single-Stage Phishing Attack

Pattern Phase	Typical Activities	Pattern Interactions
1. Research and Open Source Intelligence	• Search for open source intelligence • Establish attack objectives • Identify opportune targets	1.1 Attacker researches and strategizes about potential targets and specific objectives.
2. Planning and Preparation	• Develop attack strategy including means to avoid detection and mitigation by UIT organization • Prepare phishing attack artifacts	2.1 Attacker plans phishing attack and creates phishing artifacts (e.g., phishing email, mobile text message, phony website, malware to be implanted).
3. Phishing Operation	• Release phishing artifact via email, cellphone, rogue website, or other means • Wait for a response	3.1 Attacker initiates phishing attack through email, cellphone, rogue website, or other means.
4. Response and Information Capture	• Gain access and/or privileges to obtain greater information reach • Implant malware to achieve information objectives • Identify other opportune UIT targets and internal system information, and capture guarded and sensitive information	4.1 One or more targets unwittingly respond to phishing artifact and become a UIT. 4.2 Attacker detects or is alerted to UIT response and obtains initial information directly from UIT data entry. 4.3 Attacker implants malware on victim's machine or network. 4.4 Attacker obtains desired information via malware.
5. Attack Culmination and Exploitation	• Use captured information to directly attack UIT or UIT's organization to steal, manipulate, and/or destroy targeted assets	5.1 Attacker uses desired information in direct attack on UIT or UIT's organization to steal, manipulate, and/or destroy targeted assets.

Figure 8 shows a use case model of the single-stage attack.

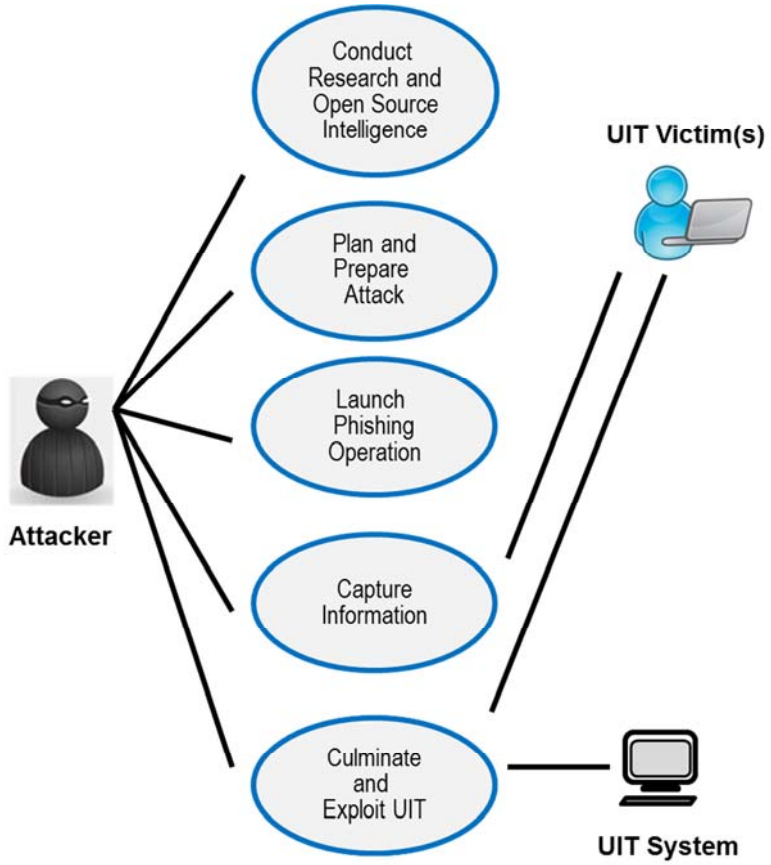


Figure 8: Use Case Model for Single-Stage Social Engineering Attack

Figure 9 shows a class model for a social engineering attack. The human participants in the Attack Participant class include the attacker and some number of UIT victims. In many phishing incidents, the attacker directs emails to a large number of potential UITs, or the potential UITs visit phishing websites. Only those who take the bait fall into the Victim subclass. The Attack Media class highlights the means that the attacker uses to obtain information, either through research in the early phases of the attack or via UIT responses, malware, or other electronic means. To carry out an exploit, attackers generate a variety of objects in the Attack Artifacts class, including email, malware, or webpages.

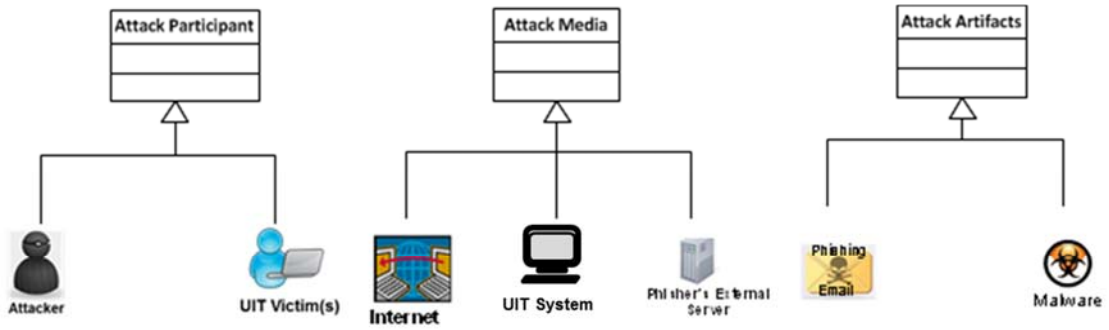


Figure 9: Attack Class Model for a Social Engineering Attack

Figure 10, a swim-lane chart, provides another perspective on the single-stage phishing attack. It shows the sequence of actions for the attacker and the victim(s).

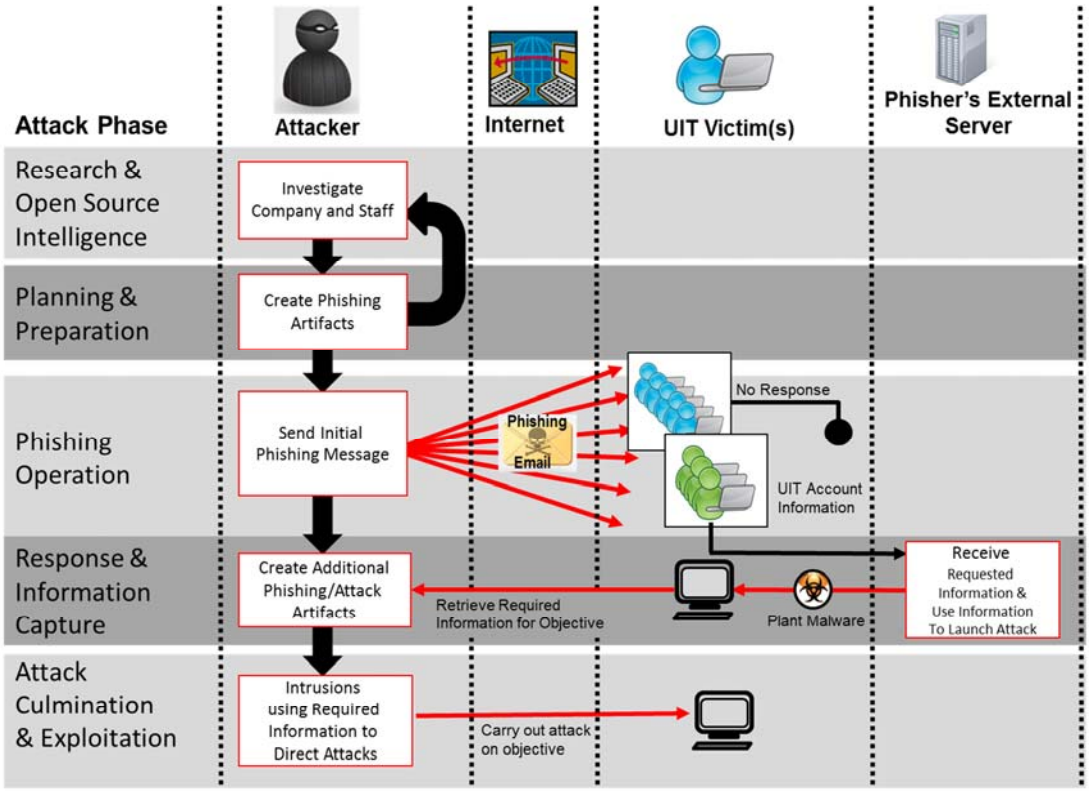


Figure 10: Swim-Lane Chart of Actions Taken by Attacker and UIT Victims in a Single-Stage Attack

Finally, the interaction view (Figure 11) peels the swim lanes apart to show each interaction and the exchanges that occur to carry out an attack. This view illustrates the collaborations of each element of the swim-lane view for a single-phase attack. The sequence of interactions shows the information exchanges during each phase of the attack.

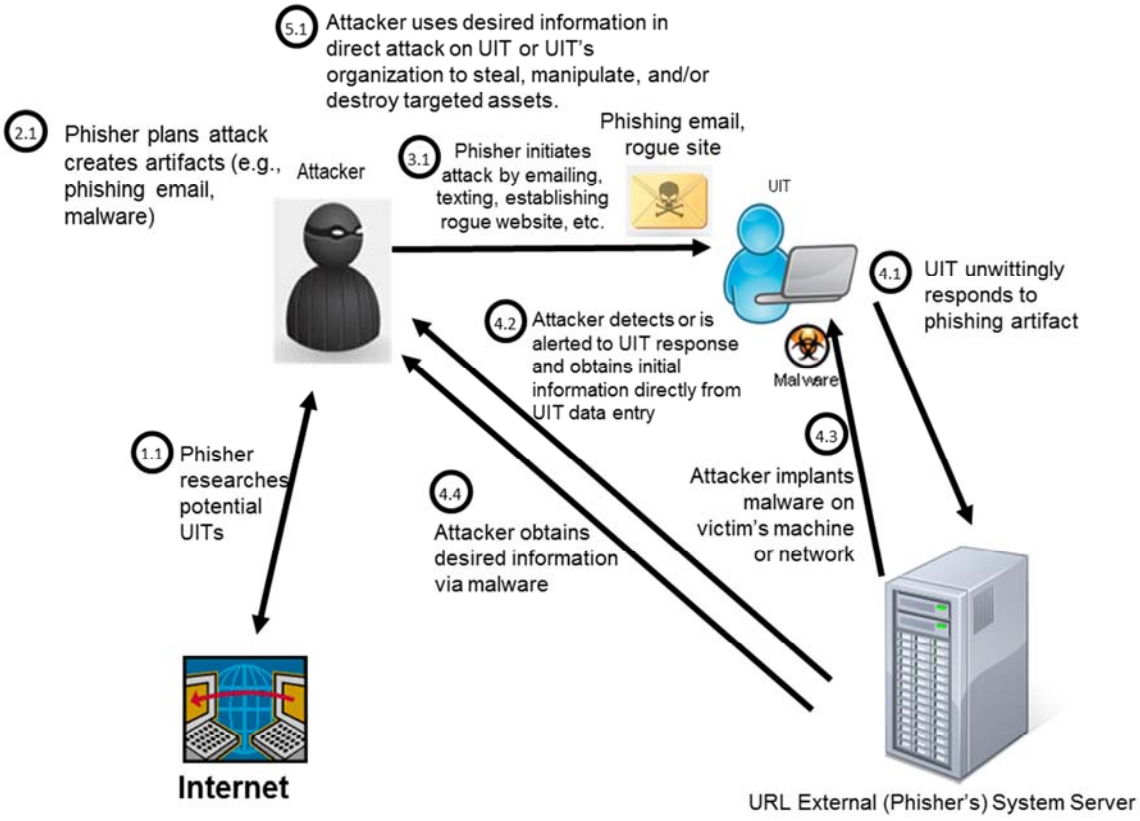


Figure 11: Interaction View Showing Object Collaboration in a Single-Stage Social Engineering Attack

6.2.2 Multiple-Stage Phishing Attack

The multiple-stage attack follows a similar pattern, but once the attacker has UIT system access, the attacker identifies other potential UITs and subsequently directs social engineering at them. The attacker may also use the access gained to probe the UIT's system to obtain various forms of internal system information. The workflow diagram in Figure 12 shows the general attack chain. This diagram identifies the ordering and decision processes involved in each phase of the exploit.

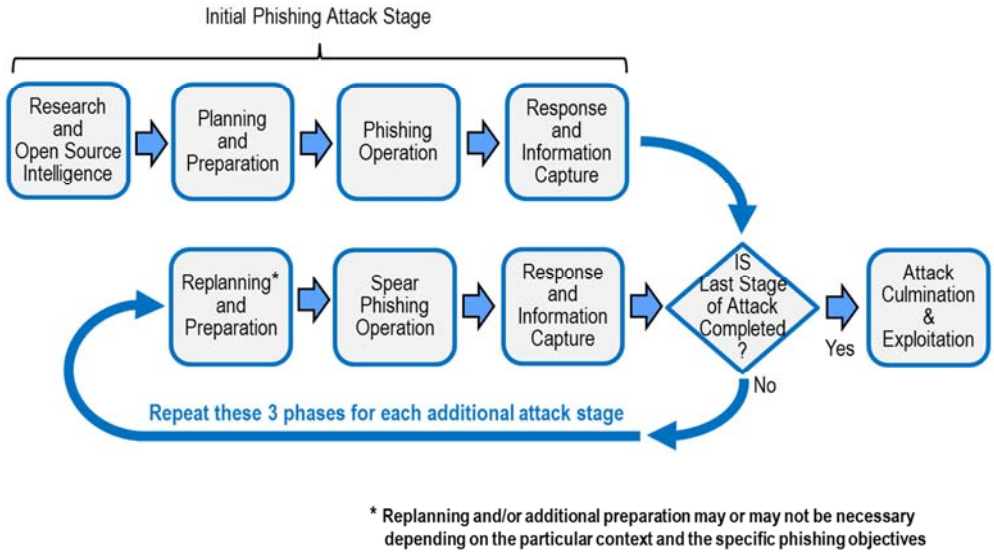


Figure 12: Workflow Diagram Attack Chain for Multiple-Stage Phishing Exploit

Table 4 shows the steps or phases of a multiple-stage phishing attack. Steps 1–4 of the single-stage attack (Table 3) still occur, but the multiple-stage attack includes additional iterative steps (shown in boldface type in Table 4) that represent the repeated planning and preparation, phishing, spear phishing, and response and information capture operations prior to conducting the delayed attack.

Table 4: Steps in a Multiple-Stage Phishing Attack

Pattern Phase	Typical Activities	Pattern Interactions
1. Research and Open Source Intelligence	• Search for open source intelligence • Establish attack objectives • Identify opportune targets	1.1 Attacker researches and strategizes about potential targets and specific objectives.
2. Planning and Preparation	• Develop attack strategy including means to avoid detection and mitigation by UIT organization • Prepare phishing attack artifacts	2.1 Attacker plans phishing attack and creates phishing artifacts (e.g., phishing email, mobile text message, phony website, malware to be implanted).
3. Phishing Operation	• Release phishing artifact via email, cellphone, rogue website, or other means • Wait for a response	3.1 Attacker initiates phishing attack through email, cellphone, rogue website, or other means.
4. Response and Information Capture	• Gain access and/or privileges to obtain greater information reach • Implant malware to achieve information objectives • Identify other opportune UIT targets and internal system information, and capture guarded and sensitive information	4.1 One or more targets unwittingly respond to phishing artifact and become a UIT. 4.2 Attacker detects or is alerted to UIT response and obtains initial information directly from UIT data entry. 4.3 Attacker implants malware on victim's machine or network. 4.4 Attacker obtains desired information via malware.
5. Replanning and Preparation	• **Replan attack strategy including means to avoid detection and mitigation by UIT organization** • **Prepare spear phishing attack artifacts**	**5.1 Attacker uses information capture in Step 4 above to replan follow-on steps for spear phishing attack. This may entail creation of new artifacts or specific attack approaches.**
6. Spear Phishing Operation	• **Execute spear-phishing** • **Wait for a response**	**6.1 Attacker initiates spear phishing attack.**
7. Response and Information Capture	• **Gain access and/or privileges to obtain greater information reach** • **Exploit more specific insider targets: financial system, secure systems, etc.**	**7.1 One or more high-value targets unwittingly responds to the spear phishing artifact and becomes a UIT.** **7.2 Phisher detects or is alerted to UIT response and obtains desired information directly from UIT data entry.**
8. Attack Culmination and Exploitation	• **Use captured information to directly attack UIT or UIT's organization to steal, manipulate, and/or destroy targeted assets**	**8.1 Attacker uses desired information in direct attack on UIT or UIT's organization to steal, manipulate, and/or destroy targeted assets.**

Figure 13 shows the use case model for this multiple-stage social engineering attack. It illustrates the initial phishing attack (left side of figure) and an additional attack (right side of figure).

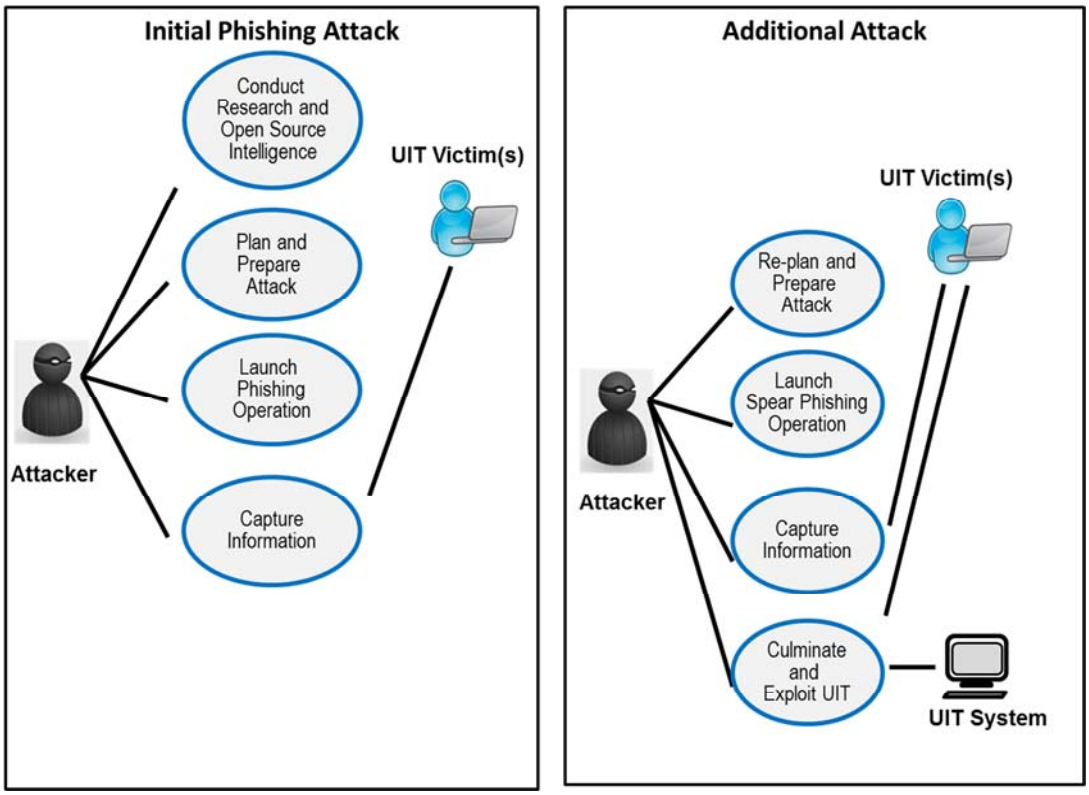


Figure 13: Use Case Model of a Multiple-Stage Social Engineering Attack

Figure 14 depicts the interaction view of object collaboration in the multiple-stage social engineering attack. The objects are derived from the single-stage attack's class model (Figure 9), but there are more instances of them.

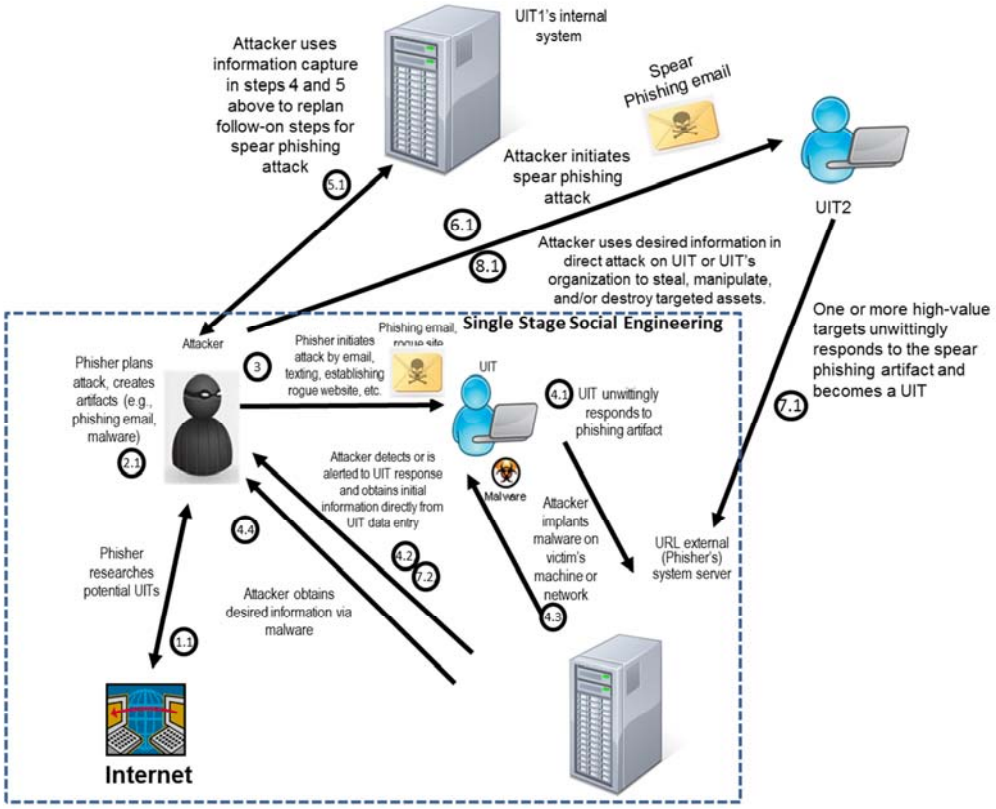


Figure 14: Interaction View Showing Object Collaboration in a Multiple-Stage Social Engineering Attack

Figure 15 illustrates Case ID #15 (Example 2 in Section 5.1.1), which represents a single-stage social engineering pattern. Figure 16 illustrates Case ID #5 (Example 1 in Section 5.1.2), which represents a multiple-stage social engineering pattern.

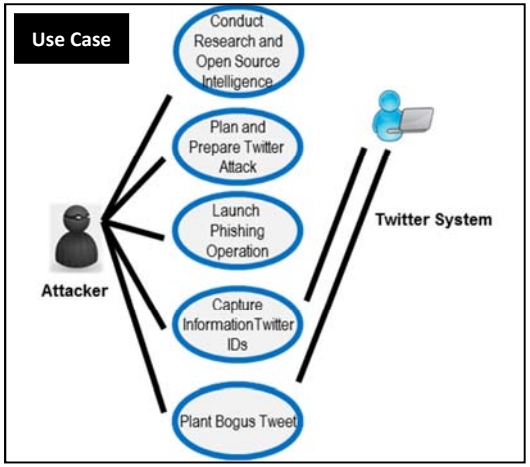
Example 2 Case #15

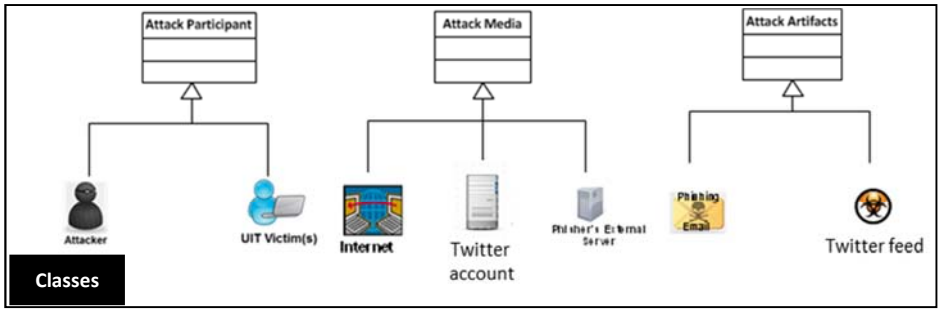
The attack is launched against a major news organization. A bogus news story is planted on the organization's twitter feed to disrupt financial markets.

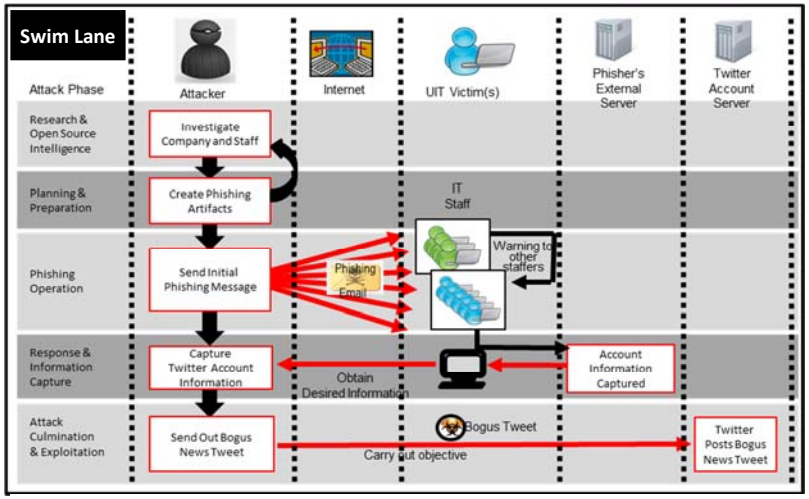
Shown here are the use case (right), participant classes (upper right), swim-lane activities (below), and collaboration model (far right).

Instead of malware, the attacker plants a bogus newsfeed.

The final use case is extended to plant the newsfeed.







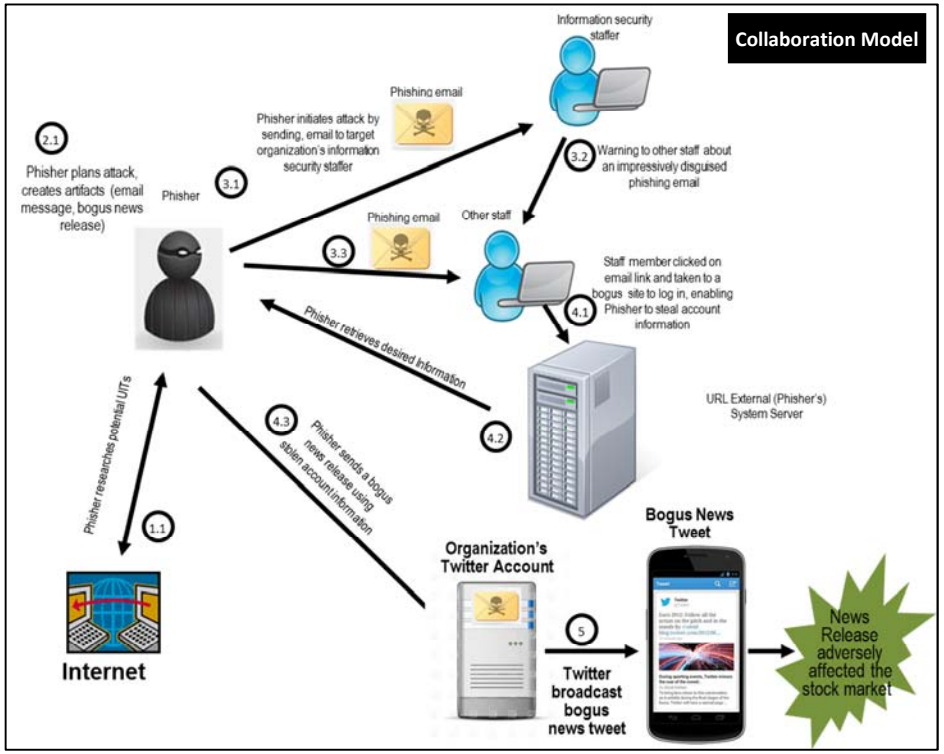


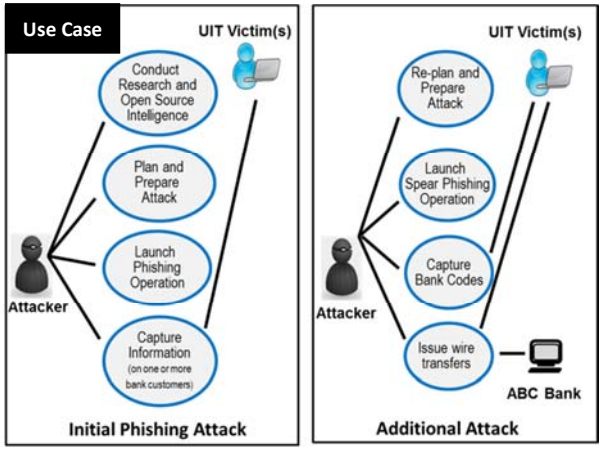
Figure 15: Illustration of Concepts and Patterns Applied to Case #15
(See Example 2, Section 5.1.1)

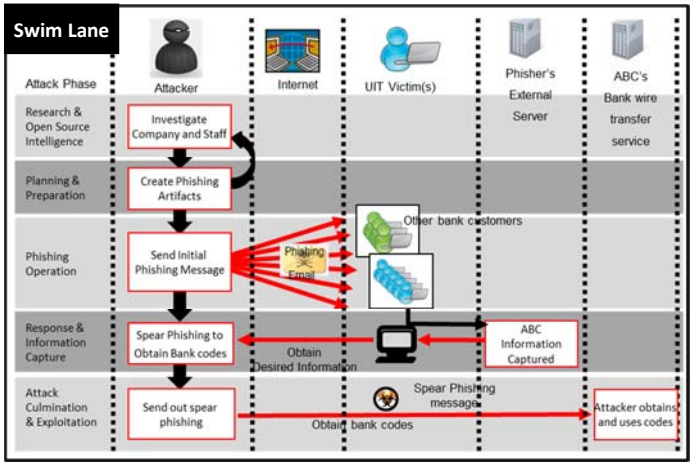
In the collaboration model (above), the attack uses a phishing message (Steps 3.1 and 3.3) against the organization's staff to obtain account information (Step 4.1). With the account information, the attacker is able to impersonate staff and plant the bogus newsfeed (Step 4.3).

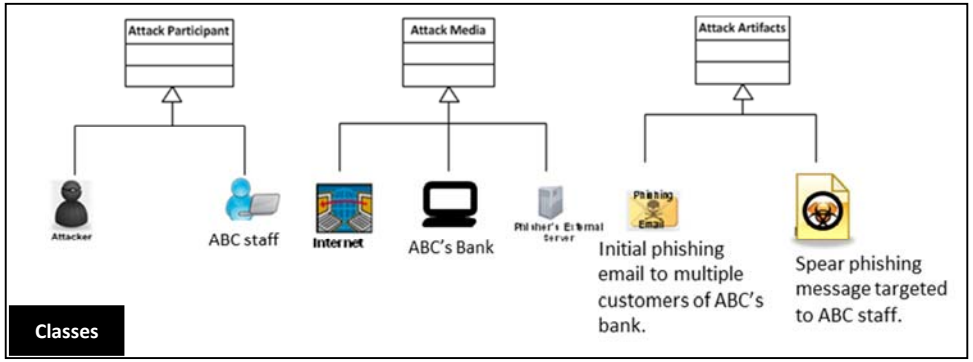
Example 1: Case #5

The attack is launched against companies using wire transfers. The target is a funds transfer account with any company using wire transfers with a specific bank. The attacker phishes the bank's customers, and ABC responds. The spear phishing goes to ABC with the intent of capturing ABC's wire-transfer authorization codes, using them to perform a series of transfers from ABC's account to accounts held by the attacker in off-shore banks.

Shown here are use cases, participant classes, and swim-lane activities. The second stage is not completely illustrated in the swim-lane graphic. After an initial set of phishing messages and receipt of a response from ABC, the attacker sends a spear phishing message to ABC to obtain wire transfer codes.







In the collaboration model, the attack uses a phishing message (#3) against bank customers. When ABC responds (#4.1, #4.2), it becomes the target of spear phishing to obtain the codes (#6.1, #7.1). With the account information, the attacker is able to impersonate ABC and perform the transfers (#8.1).

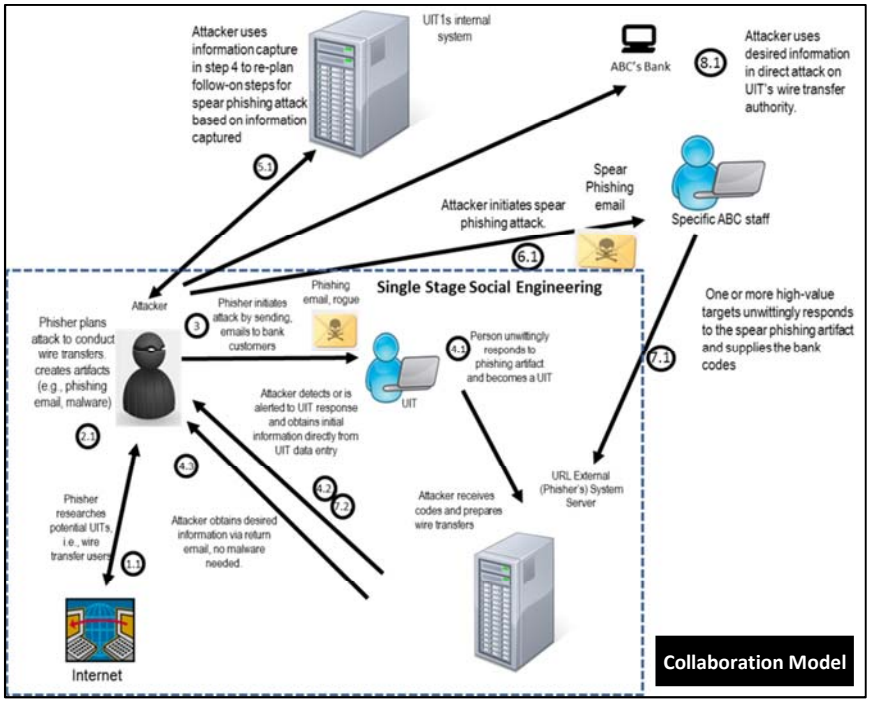


Figure 16: Illustration of Concepts and Patterns Applied to Case #5

(See Example 1, Section 5.1.2)

6.3 Descriptive System Dynamics Model

Another way to describe and characterize social engineering exploits is to use system dynamics modeling, which helps analysts model and analyze critical behavior within complex socio-technical domains as it evolves over time [Sterman 2000]. Here we describe a system dynamics model that captures the complex interactions within a social engineering scenario. The model focuses on key aspects of the social engineering UIT incident; in a later subsection, we refine the model to illustrate leverage points for mitigation.

6.3.1 Causal Loop Diagrams

Figure 17 summarizes the notation used in this section. Causal loop diagrams show qualitatively how related variables affect each other [Meadows 2008]. The nodes indicate variables, and the connecting arrows show the relationships between them. Arrows are labeled to indicate how the variable at the arrow's source influences the variable at the arrow's target. Basically, a positive influence indicates that the values of the variables move in the same direction and so is labeled with an "S," whereas a negative influence indicates that they move in the opposite direction, indicated by an "O" label.

A connected group of variables can create a path that is referred to as a feedback loop. The type of feedback loop is determined by counting the number of negative influences along the path of the loop. An odd number of negative influences indicates a balancing loop; an even (or zero) number of negative influences indicates a reinforcing loop. Balancing loops often represent actions that an organization takes to mitigate (or control) a problem. Reinforcing loops often represent the escalation of problems but may include problem mitigation behaviors.

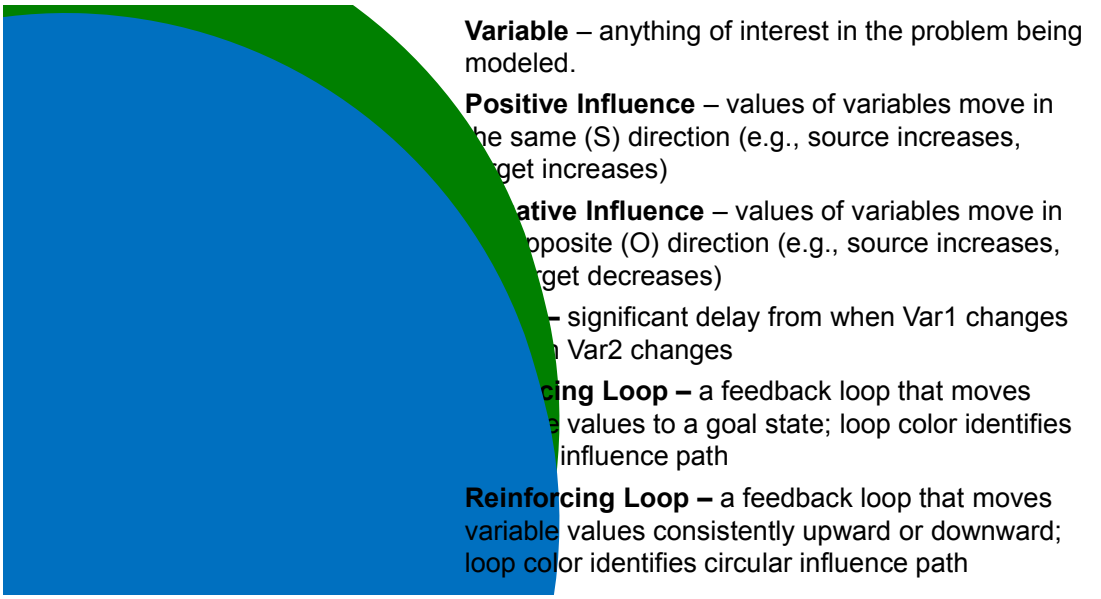


Figure 17: System Dynamics Notation Used in Abstract Models

6.3.2 Confirmatory Bias Loop

A feedback loop relevant to the social engineering problem is a reinforcing loop called the Confirmatory Bias Loop, shown in Figure 18. Confirmatory bias is the tendency of decision makers to pay attention to data that supports (or is at least consistent with) their past decisions and to downplay conflicting information [Sastry 1989, Staw 1989]. This bias can skew the basis for decision making so that alternate decisions are overlooked in favor of the preferred decision. This form of the model illustrates the impact of cognitive limitations (attention limits, cognitive biases, errors in judgment, and decision making) that were discussed in Section 4.3. The associated feedback loop portrays the reinforcing nature of a decision maker's cognitive process.

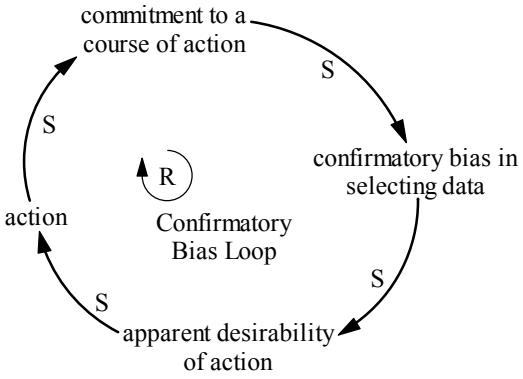


Figure 18: Confirmatory Bias

6.3.3 Phishing Exploits in Social Engineering

Figure 19 shows at a high level how phishing exploits often unfold in social engineering exploits. Feedback loop B1 (shown in purple) shows the initial stage of the attack where an outsider desires a certain level of access privilege (shown at the bottom of the figure) to carry out an exploit. Research and open source intelligence gathering provide some knowledge of the insider to plan and prepare an effective phishing exploit. As indicated in current literature, visual deception and obfuscation can increase the chances that the insider will fall for the deception and provide the outsider with the information to increase the outsider's access privilege. If greater privilege is needed, the outsider may use the gained knowledge to deepen the access through even more narrowly targeted spear phishing exploits, as shown in the reinforcing feedback loop R1 (in red). Such a campaign is a multiple-stage phishing attack that gradually exploits information accumulated by the outsider.

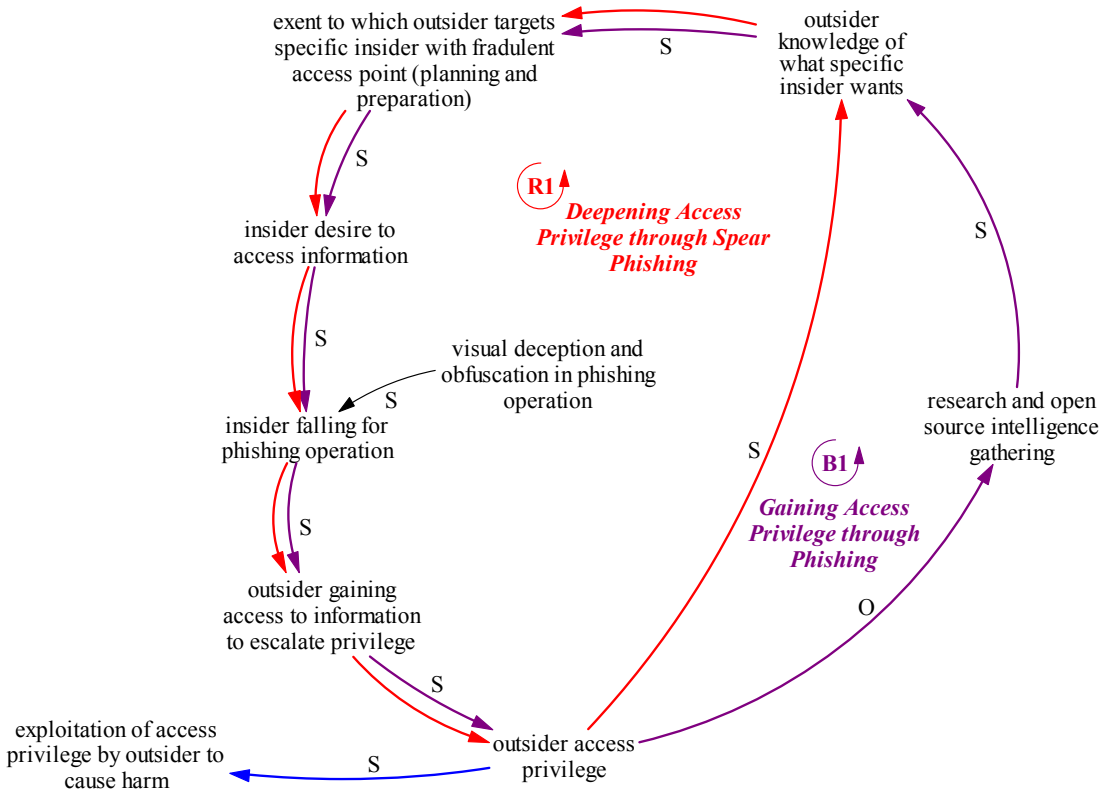


Figure 19: Causal Loop Diagram of Phishing Exploits

6.3.4 Confirmatory Bias in Social Engineering

Figure 20 shows the role confirmatory bias can play in social engineering exploits. Two situations are depicted. In the first, the insider desires access to information supplied by the outsider's created (deceptive) scenario, as depicted in the R2 (orange) feedback loop. The second is where the insider desires to be helpful to the malicious outsider in need as depicted in the R3 (green) feedback loop. Both loops portray the reinforcing of trust in the outsider's authenticity and the subsequent desire to access information or to be helpful.

The key to the confirmatory bias tendency is that the growing desire to believe the scenario put forth by the outsider leads the insider to focus on evidence that confirms the legitimacy of the scenario and ignore evidence to the contrary. The model depicted in this figure reflects research findings reviewed in Section 4.3:

- High levels of cognitive workload can increase the chances that the insider will believe the deceptive scenario painted by the outsider and trust the outsider's authenticity.
- The insider's overall awareness of the risks of social engineering also plays a role in the trust the insider places in the outsider's scenario.
- Creating a sense of urgency increases the chances of falling for the deception, either as a need to be helpful or to access the information (phishing exploit) provided by the outsider.

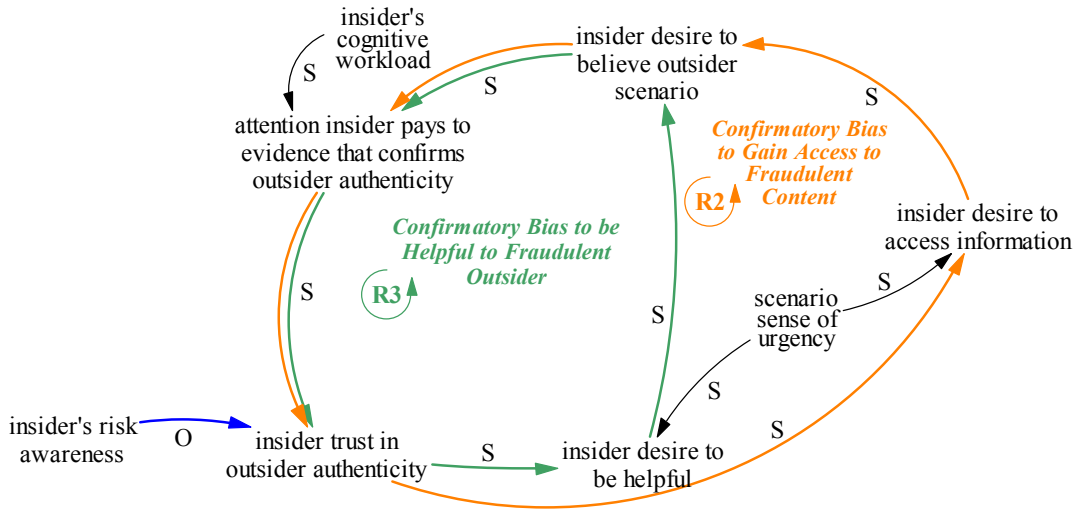


Figure 20: Confirmatory Bias in Social Engineering Exploits

6.3.5 Integrated Model of the Social Engineering Problem

Figure 21 integrates and extends the causal loop diagrams described above. The insider's desire to access information is reinforced through the outsider's planning and preparation from the R1 and B1 feedback loops, as well as the insider's confirmatory bias. Once the insider's desire to access the information provided by the outsider (loop R2) or to help the outsider (loop R3) reaches a threshold, the insider provides the outsider undeserved access, resulting in the outsider attack, as shown in the bottom part of the figure. Trust in the outsider's authenticity is enhanced by a credible scenario provided by the outsider, supported by an accurate impersonation, as shown in the upper left.

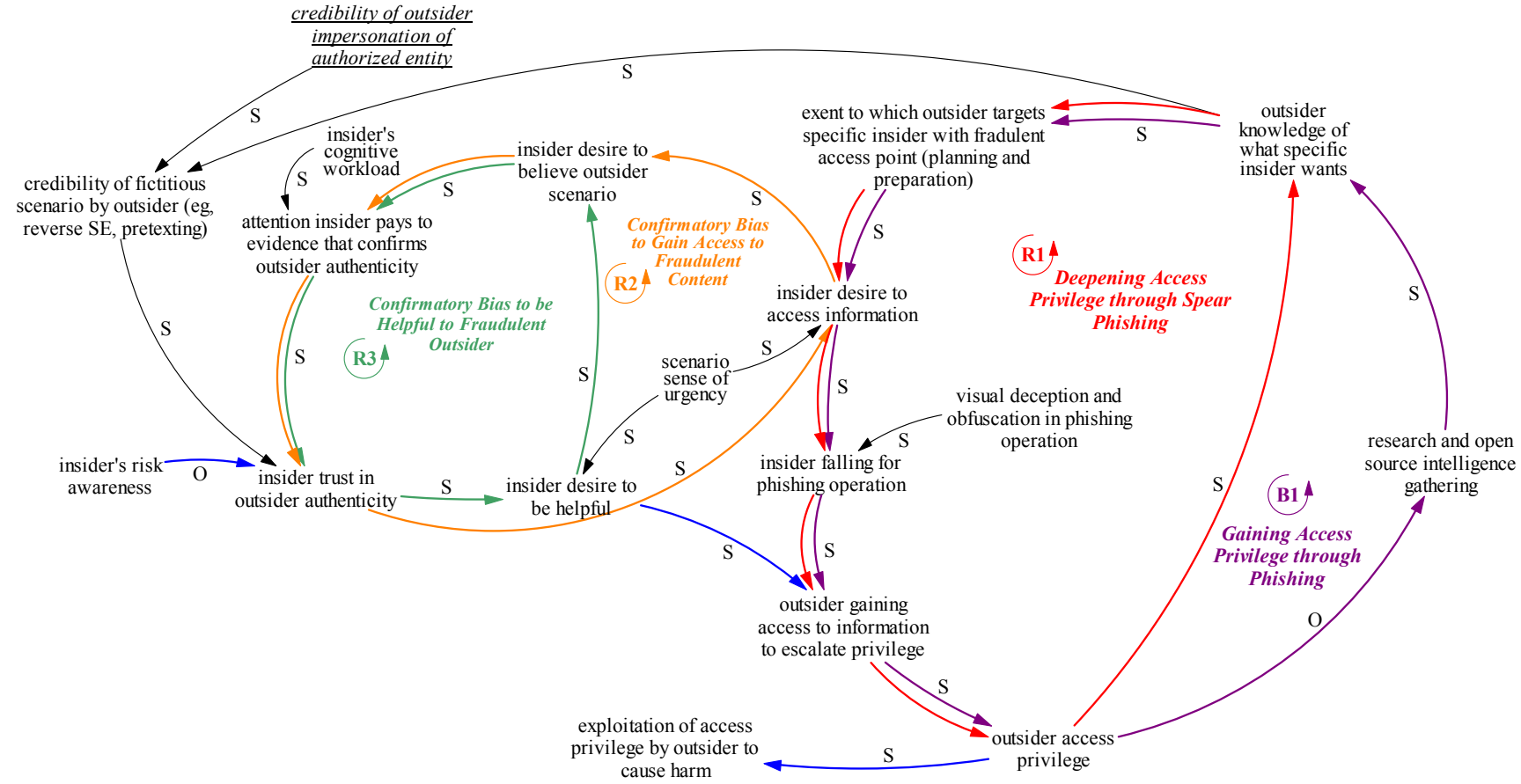


Figure 21: Causal Loop Diagram of Social Engineering of Insiders by Outsiders

6.4 Ontology of Social Engineering Tactics

6.4.1 Need for a Taxonomy

The distinctive essence of social engineering is the psychological manipulation of the human decision-making process. It is also what makes it so effective and such a hard technique to mitigate. It is essential to have a clear theoretical understanding of the manipulation process in order to classify incidents, understand their elaboration, and devise mitigation techniques. This need for a theoretical framework is especially crucial because of the central role awareness training plays in preventing social engineering attacks. Most antiphishing training, for example, focuses on details of the presentation format such as amateurish design and spelling mistakes. We suspect that training on the techniques of psychological manipulation might enhance the effectiveness of that training.

6.4.2 Social Engineering Tactics Described in Research Literature

For this project, we produced a taxonomy of social engineering tactics. While our research did not propose any new social engineering tactics, the development of our taxonomy represents a contribution that brings together disparate efforts to characterize social engineering tactics. The tactics in our taxonomy have been developed since 2001 by three principle authors. To provide attribution of these tactics to the authors who first described them, we use a labeling process that uses the first letter of the author's name plus a sequence digit.

In 2001 Granger enumerated friendliness (G1), impersonation (G2), conformity (G3), decoying (G4), diffusion of responsibility (G5), and reverse social engineering (G6) as so-called "hacker tactics" [Granger 2001]. Also in 2001, Cialdini published the first edition of *Influence: Science and Practice*, which added the three tactics of reciprocity (C1), perceptual consistency (C2), conformance (C3), trust (C4), fear (C5), and scarcity (C6), which were not on Granger's list [Cialdini 2001]. Finally, in 2006 Lena Laribee published *Development of Methodical Social Engineering Taxonomy*, which included five new tactics based on her research: sympathy (L1), guilt (L2), equivocation (L3), ignorance (L4), and affiliation (L5) [Laribee 2006a]. The full list of tactics on which we based our taxonomy is as follows:

- (G1) friendliness
- (G2) impersonation
- (G3) conformity
- (G4) decoying
- (G5) diffusion of responsibility
- (G6) reverse social engineering
- (C1) reciprocity
- (C2) perceptual consistency
- (C3) conformance
- (C4) trust
- (C5) fear (obedience to authority)
- (C6) scarcity
- (L1) sympathy
- (L2) guilt
- (L3) equivocation
- (L4) ignorance
- (L5) affiliation

One question that arises is the effectiveness of these tactics and their connection to psychological constructs. As discussed in Section 4.1.3, this was studied empirically by Workman, who found statistically significant correlations between five of the six C-factors (above) and susceptibility to social engineering exploits [Workman 2008]. Workman's experiments were based on the Elaboration Likelihood Model (ELM), which makes a distinction between central and peripheral

routes of persuasion, which we find useful in conceptualizing social engineering tactics. Workman describes the distinction as follows:

> *The ELM distinguishes "central" from "peripheral" routes of persuasion, where a central route encourages an elaborative analysis of a message's content, and a peripheral one is a form of persuasion that does not encourage elaboration (i.e., extensive cognitive analysis) of the message content. Rather, it solicits acceptance of a message based on some adjunct element, such as perceived credibility, likeability, or attractiveness of the message sender, or "a catchy" phrase or slogan [Miller 2005]. For example, celebrities are frequently used to sell products with which they have no obvious or special expertise, and consumers often purchase these products because "they think they know" and like or identify with the celebrity [Cacioppo 1986]. Peripheral route persuasion is an important element in social engineering scams because it offers a type of shield for the attacker [Mitnick 2002].*

6.4.3 Design Goals for the Taxonomy

The goals in building the taxonomy were as follows:

- *Comply with a strict class hierarchy.* The tactics described by Granger, Cialdini, and Laribee form a simple paratactic list. We believe that the tactics actually form a conceptual hierarchy that should be elucidated. Therefore, our objective was to represent the tactics using a class hierarchy.
- *Clarify the nomenclature.* It is important to be precise about the terms used. For example, because "friendship" is not only a deception technique, we replaced the label for this tactic with "false friendship."
- *Be as comprehensive and as fine-grained as possible.* Our taxonomy differs from those of our sources in a number of ways:
  - At the root of the tree, we included two branches for nondeceptive forms of social engineering: coercion and inducement. These fit the general definition of social engineering, but do not fit the more specific definition of UIT social engineering: that is, they do not depend on deceiving the target of the engineering; coercion and inducement would more likely fall into the broader taxonomy of insider threats (including malicious insider threat). In the last decade or so the concept of social engineering has been overwhelmed by the prevalence of cyber social engineering, which is almost always based on deception.
  - We have introduced some subcategories. We distinguish between the true psychological mechanisms, such as false friendship, and some rhetorical devices, such as distraction and ambiguity, that increase the probability of the psychological mechanisms succeeding.
  - We have broken many of the tactics down into subtactics. Thinking in terms of a formal class hierarchy encourages making those fine-grained distinctions; it will make the ontology more usable when it becomes encoded in the Web Ontology Language (OWL).
- *Tie to vulnerabilities and mitigation techniques.* Most of the tactics have a corresponding psychological characteristic that is exploited and a corresponding mitigation technique that blocks the activation of that characteristic.

6.4.4 The Taxonomy

The taxonomy in its current form is shown below. Each level of the hierarchy is related to its parent by an "is-a" relationship. For example, "flirtation" is a form of false friendliness, which is a psychological mechanism for deceit, which is a social engineering tactic. Throughout, the class names are in the singular to facilitate importation into the OWL.

- Social engineering tactic
  - Deception
    - Psychological mechanism
      - False friendliness (C4, G1)
        - Attractiveness
        - Likability
        - Cordiality
        - Convincingness
        - Flattery
        - Flirtation
        - Pleasant language
          - Spoken language
            - Greetings ("Thank you")
          - Body language
            - Smiles
      - False sympathy (L1)
        - Sharing unhappiness
        - Sharing suffering
        - Displaying concern
        - Displaying desire to alleviate negative feelings
      - False authority (C5)
        - Displaying knowledge
        - Displaying wisdom
        - Displaying power
        - Exploiting sense of duty
        - Self-promotion
          - Establishing affiliation (L5)
            - Name dropping
            - Suggesting membership in inner circle
          - Reducing suspicion of attacker's motives
      - Impersonation (G2)
        - Assertion of identity without study
        - Assumption of identity after study
        - Imitation
          - Voice disguises
          - Speech patterns
          - Badges
        - Local knowledge
          - Organizational chart
      - Conformity

- Providing social proof (G3, C3)
- Guilt
  - Claiming victim has obligation to help (C1)
  - Convincing victim to take responsibility for misfortune (L2)
- Commitment (C2)
  - Securing initial commitment
  - Leveraging initial commitment

- Rhetorical technique
  - Diffusion of responsibility (G5)
  - Decoying (G4)
    - Distraction
    - Change of focus
    - Reciprocity rule
  - Equivocation (L3)
    - Double meanings
      - Ambiguity
      - Innuendo
      - Creating uncertainty
    - Semantic shift
  - Feigned ignorance (L4)
    - Pretending to be uninformed
    - Extracting information to help “new employee”
  - Scarcity (C6)
    - Creating sense of urgency
    - Pressuring victim for information
- Complex game plan
  - Reverse social engineering (G6)
    - Sabotage
    - Advertising
    - Assisting

- Coercion (punishment)
  - Blackmail
  - Intimidation, bullying
- Inducement (Reward)
  - Bribery
  - Collusion

6.5 Implications for Mitigation of Social Engineering Exploits

While it is beyond the scope of this work to examine current mitigation practices, the foregoing discussion and characterization of social engineering attacks in terms of possible contributing factors (especially organizational and human factors) and patterns help to inform a brief consideration of possible mitigation approaches and strategies. Here we briefly discuss and speculate about possible implications for mitigations suggested by systematic analyses of patterns and models of social engineering exploits discussed in previous sections of this report.

6.5.1 Implications of Patterns and Characterizations

As we observed in our presentation of patterns in UIT social engineering exploits, the kill-chain pattern indicates that the adversary must progress successfully through each phase of the chain before the desired objective can be achieved: the chain and the adversary can be disrupted by a single mitigation that is applied successfully to one phase of the chain. This observation has strong implications for concepts for and approaches to mitigation. In the present context, a sophisticated multiple-stage social engineering attack (such as one involving phishing followed by spear phishing) aims to breach successive layers of organizational defenses by progressively gaining access through social engineering methods. The attack continues iteratively, and sometimes opportunistically, to take advantage of individual or organizational responses until the final layer of defense is breached. Because the ultimate success of a multiple-stage attack depends on the success of each individual (i.e., iterative) stage leading up to the final attack, the kill-chain approach affords a UIT organization multiple opportunities to detect and defeat such attacks.

A systematic analysis of patterns in workflow diagrams or use case representations can reveal points at which opportunities for mitigation arise, leading to possible mitigation approaches. For example, using the workflow and kill-chain pattern for a single-stage social engineering attack, we may identify the following types of mitigation that would apply to different phases of the attack:

- *Research and Open Source Intelligence phase*—To combat efforts of the social engineer to acquire information to exploit about a company or its employees, some steps may be taken to limit that amount of information or details that might be exploited. It is not possible or desirable to completely eliminate this type of information, but the organization may benefit from instilling some controls and safeguards in its public relations and information dissemination processes to avoid excessive disclosures of such information. Similarly, employees may be given direction or policies about the type of information to avoid making public through social media sites.
- *Planning and Preparation phase*—Less potential exists for organizations to impact the attacker's planning and preparation for the attack. However, efforts should be made to make it difficult or expensive to copy organizational artifacts that make a spoofing email or website look legitimate. This could impair or discourage attacker's efforts to masquerade or impersonate organizational assets. Anticounterfeiting strategies such as encrypted email emails are well known but not commonly used.
- *Launch Operation phase*—Phishing exploits target human psychological characteristics and limitations, so improved training and awareness are an organization's most potent mitigation tools. Periodic injection testing and associated training may be used to maintain staff vigilance and knowledge about the most current social engineering tactics. Organizations also should strive to maintain productive work attitudes and information security awareness through human factors and organizational practices. Effective management and resource planning can help ensure employee productivity and avoid stressful work environments that may lead to errors in judgment.
- *Information Capture phase*—Organizations should enable and maintain improved tools for computer and network defense cyber monitoring to keep up with the rapidly evolving kinds of exploits that adversaries use. Cybersecurity systems that locate malware and other threats

include antivirus, data loss protection (DLP) tools, and security information and event management (SIEM) products.

- *Culmination and Exploitation phase*—In this phase, mitigations are the same as for the Information Capture phase.

Figure 22 illustrates these strategies.

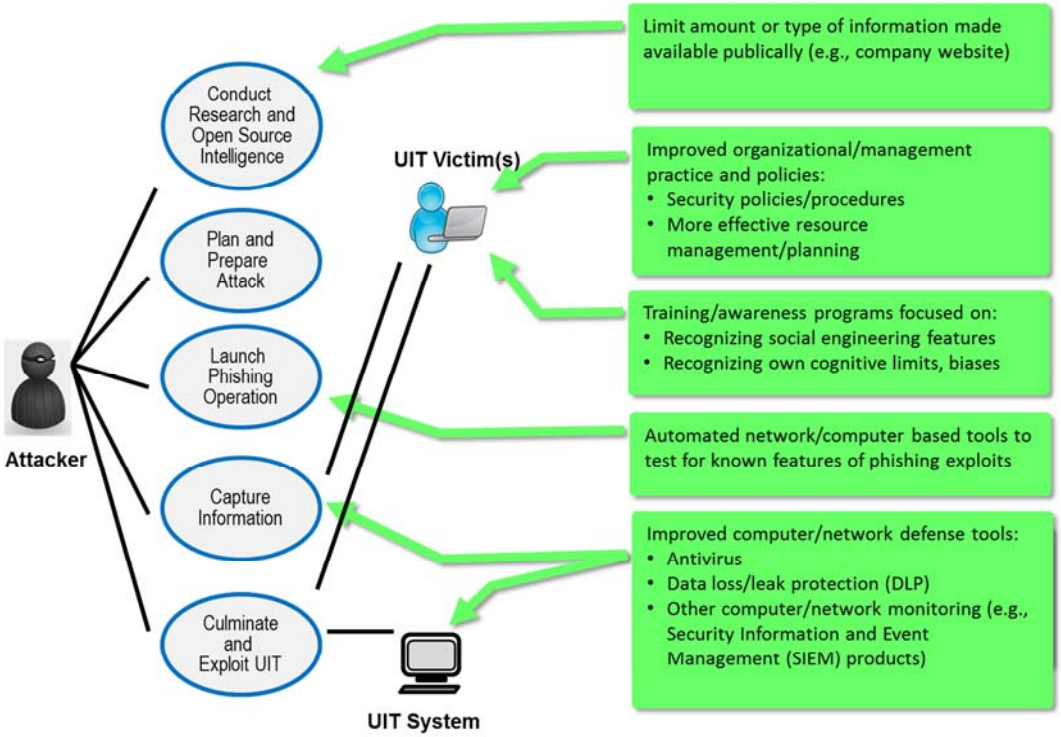


Figure 22: Mitigation Strategies that Apply to Different Phases of an Attack

6.5.2 Implications of Social Engineering Tactics Ontology

As noted, most of the tactics exploit a particular psychological characteristic, which suggests a corresponding mitigation technique to block the exploitation. For example, the Feigned Ignorance tactic works because people have a natural tendency to want to help others, compounded by a desire to show off their own expertise. The mitigation technique is to be more mindful and skeptical when assessing claims of cluelessness. Eventually the vulnerabilities and mitigation techniques will be integrated into a true social engineering ontology, but for the present purposes, consider conceptual associations between some of the tactics, vulnerabilities, and mitigations, as shown in Table 5.

Table 5: Social Engineering Tactics, Vulnerabilities, and Mitigations

Tactic	Vulnerability	Mitigation Question
False friendliness	Desire to trust	*Why is this guy being so nice to me?*
False sympathy	Desire to share sorrows	*Does this guy really care about my child's accident?*
False authority	Fear of power	*How can I be sure this guy is who he says he is?*
Impersonation	Assumption that things are as they seem	*Is this guy's uniform a little too perfect?*
Conformity	Desire to fit in	*Isn't a little surprising that everyone else is doing this?*
Scarcity	Tendency to value scarce things	*Wait a second—does this really have to be done right now?*

A more complete analysis of the relationships between social engineering tactics and psychological vulnerabilities will yield a more comprehensive list of mitigation questions. Such considerations may form the basis of much more detailed training that goes beyond cybersecurity awareness of possible social engineering exploits. More targeted training should focus on challenging the trainees to recognize these psychological tactics and their own psychological vulnerabilities, in order to prepare individuals to be more vigilant. Generating a list of mitigation questions, as illustrated in Table 5, can support more detailed, role-based training approaches to countering social engineering attacks.

6.5.3 Implications of System Dynamics Model

The system dynamics model may be used to help identify possible mitigation strategies. This section applies system dynamics modeling to a hypothetical case, depicted in Figure 23. Because this example does not reflect all possible approaches that adversaries might take in executing the social engineering attack, implications drawn from the example are not exhaustive. However, they are instructive and representative of how the analysis can reveal opportunities for applying measures to circumvent (balance) the actions of malicious attackers.

Figure 23 shows how the reinforcing feedback loops (R1, R2, and R3) involving the escalation of the phishing exploit and the cognitive limitations of the insider can be dealt with by balancing feedback loops:

- Feedback loop B2 (light blue) represents organizational processes aimed at reducing the effectiveness of social engineering exploits in taking advantage of insiders. Feedback loop B2 involves the recognition of the exploitation by the organization and improved training on the nature and risks of social engineering to organizational insiders. Specifically, the organization provides more effective training and awareness about how malicious outsiders use obfuscation and social engineering techniques to deceive insiders. Such training may involve various topics relevant to human factors described in Section 4, aimed at raising self-awareness about cognitive limitations and biases, fostering greater security awareness and more accurate risk perception, and encouraging more diligent application of computer security policies.
- Feedback loop B3 (dark blue) represents organizational processes aimed at reducing the effectiveness of early-stage social engineering activities that aim to acquire intelligence about the organization that may be used in an initial phishing attack. Specifically, the mitigation

approach seeks to reduce the amount of publicly available information about the organization and its employees that malicious outside social engineers can use to develop initial attack plans and associated artifacts for luring insiders into their traps.

Not shown in the figure are other possible opportunities for mitigation that would be aimed at different parts of the system dynamics model. For example, mitigation in the form of more effective firewalls or automated tools for recognizing flaws in phishing emails might be applied to balance spear phishing efforts in R1.

Of course, all of these mitigation approaches are hampered by time delays. The longer the delay associated with the organization's mitigation action, the less effective it will be in preventing the successful execution of social engineering exploits.

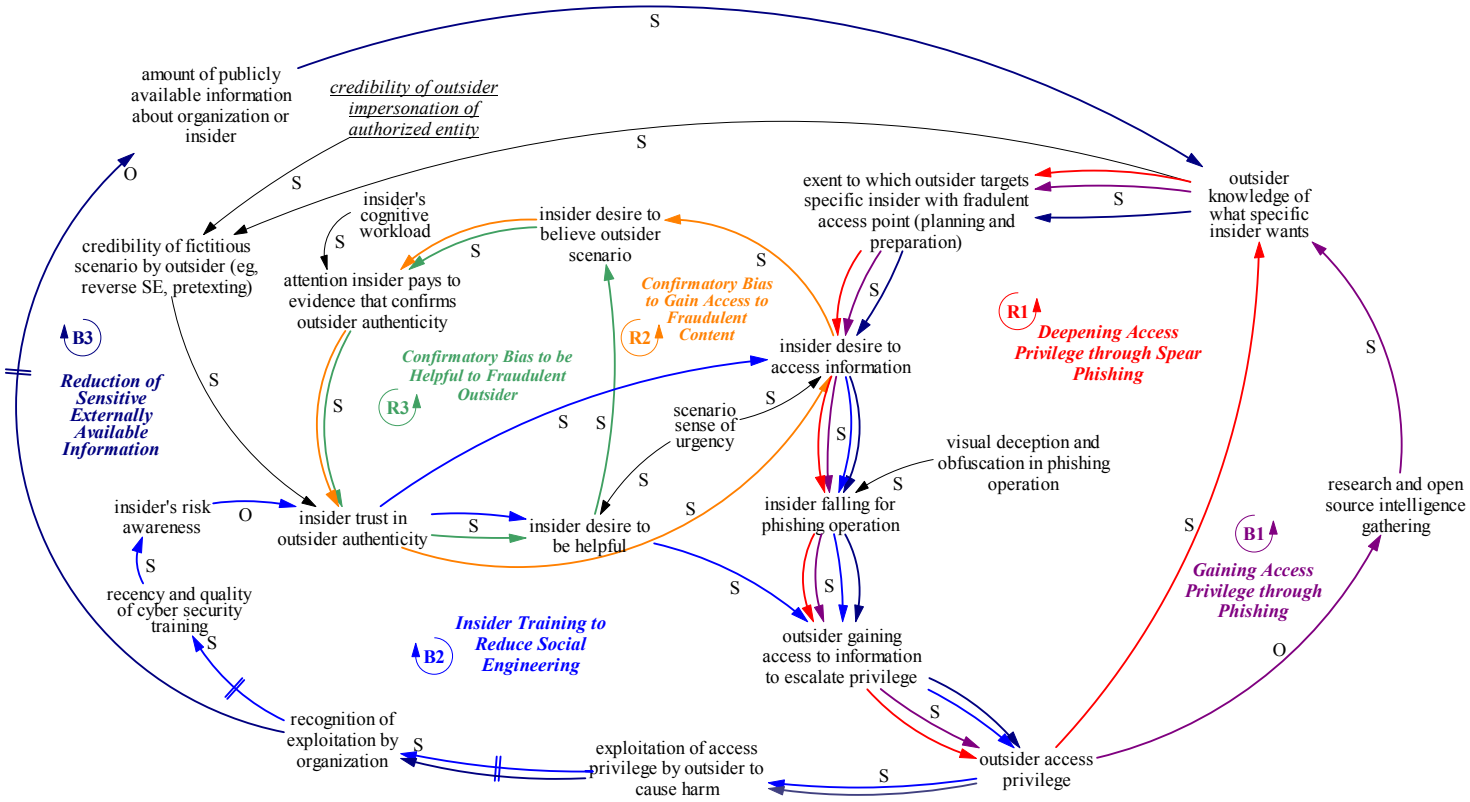


Figure 23: Causal Loop Diagram of Avenues for Social Engineering Mitigation

6.5.4 Summary and Conclusions About Mitigation

Our analysis of collected case studies reveals a number of commonalities or patterns that may be useful to take into account when developing mitigation tools or strategies:

- Using kill-chain concepts, we recognize common features, or building blocks, in single-stage and multiple-stage social engineering attacks. Each stage contains recognizable patterns or building blocks that compose the attack. Each stage includes multiple phases. To be successful, the attack must succeed at every phase.
- Some phases represent actions of the attacker, while other phases represent actions of UIT victims. Mitigation strategies and tools should be crafted to target specific characteristics in each attack phase.

Our review and analysis of research and case studies suggest the following mitigation strategies to reduce the effectiveness of social engineering attacks:

1. Organizations should examine their management practices to ensure that they meet human factors standards that foster effective work environments to minimize stress (e.g., minimizing time pressure and optimizing workload) and encourage a healthy security culture. Because employees may perceive information security compliance as interfering with job functions, it is important for organizations to allocate a certain amount of employees' time to fulfilling the compliance requirements.
2. Organizations should develop and deploy effective staff training and awareness programs aimed at educating users about social engineering scams, including learning objectives to help staff notice phishing cues, identify deceptive practices, and recognize suspicious patterns of social engineering exploits.
3. Research is required to develop more effective network and workstation monitoring tools to recognize attributes of social engineering artifacts (e.g., emails).
4. The research and stakeholder community should develop mitigations that apply to specific attack phases, such as the following:
   - *Research and Open Source Intelligence phase*—Both the organization and individual employees may benefit from limiting the amount of information available on organizational websites or individuals' social media sites, which might be exploited by outsiders.
   - *Planning and Preparation phase*—Efforts should be made to make it difficult or expensive to copy organizational artifacts that make a spoofing email or website look legitimate. Anticounterfeiting strategies that allow encrypted emails are well known but not commonly used.
   - *Launch Operation phase*—Phishing exploits target human psychological characteristics and limitations, so improved training and awareness are an organization's most potent mitigation tools. Periodic injection testing and associated training may maintain staff vigilance and knowledge about the most current social engineering tactics.
   - *Information Capture and Culmination and Exploitation phases*—Organizations should enable and maintain improved tools for computer and network defense cyber monitoring to keep up with the rapidly evolving kinds of exploits that adversaries use.

7 Conclusions

As we have noted, more than 40% of security professionals report that their greatest security concern is employees accidentally jeopardizing security through data leaks or similar errors [AlgoSec 2013]. The previous report by the CERT Insider Threat team provided an initial examination of this problem [CERT 2013]. That report characterized the UIT by developing an operational definition, reviewing relevant research to gain a better understanding of its possible causes and contributing factors, and providing examples of UIT cases and the frequencies of UIT occurrences across several categories. That report also documented our first design of a UIT feature model, which captures important elements of UIT incidents. As a follow-on to that study, this report seeks to advance our understanding of UIT contributing factors by focusing on a particular type of UIT incident, social engineering.

Social engineering is a key component of one of the four previously identified threat vectors of UIT incidents, UIT-HACK, which was defined as electronic entry by an outside party, such as malware and spyware. This threat vector allows attackers to use an organization's unwitting insiders to help the attacker achieve outcomes that put the organization and its assets at risk. What is unique about this threat vector is that the malicious party exploits multiple human, organizational, and demographic factors to deceive the UIT victim into unwitting actions that support or advance the social engineering attack. Our literature review identified empirical and survey-based studies that suggest potential human, organizational, and demographic factors that contribute to the success of social engineering attacks; however, no behavioral research to date has provided definitive evidence of causal factors and associated mitigation strategies. It is reasonable to conclude that there are many contributing factors, including some inferred human and organizational factors, and it is not possible to identify any single cause of (or an associated mitigation approach to) social engineering UIT attacks.

To complement the current literature review, our team collected cases of UIT incidents, with a specific focus on social engineering. We analyzed the cases to determine the extent to which case documentation could be related to findings in the literature so that common concepts or approaches could be identified. Further, we examined the cases for any patterns that relate to methods, targets, contributing factors, and progression of the attacks.

7.1 Overview of Findings

Our findings are summarized as follows:

- There is, at best, a weak association between social engineering susceptibility and various demographic factors, emphasizing the need for more research to clarify or disambiguate certain relationships. Possible age effects may be confounded with related factors such as amount of experience. Research suggests that it may be possible to use personality factors to identify higher risk individuals or tailor training topics more directly to individuals with vulnerable personality factors, but further research is necessary. There have been no formal studies of possible cultural differences, but research to date does not indicate any cultural differences in response rates to social engineering scams.
- Organizational factors can produce system vulnerabilities that adversaries may exploit in social engineering attacks. Management systems or practices that provide insufficient training, inadequate security systems and procedures, or insufficient resources to successfully complete tasks may promote

confusion, reduce understanding, and increase employee stress, all of which increase the likelihood of errors or lapses in judgment that enable the attacker to successfully breach defenses.

- Academic research has identified possible human factors involved in UIT social engineering susceptibility: the research suggests that relevant human factors include insufficient attention or knowledge that would enable users to recognize cues in socially engineered messages; cognitive biases or information processing limitations that may lead the UIT victim to succumb to deceptive practices and obfuscation; and attitudes that ignore or discount risks, or that lead individuals to take shortcuts around information security compliance they feel is interfering with job functions. However, there has been little or no empirical validation in the UIT context, nor has there been any relevant reporting of these factors in real-world case studies. This situation has hampered validation of empirical findings in actual case studies.
- Analysis and conceptual modeling of collected case studies reveal a number of commonalities or patterns that may inform the development of mitigation tools or strategies. Social engineering attacks may be characterized as comprising a single stage or multiple stages, and within each stage there are recognizable patterns or building blocks that compose the attack.

7.2 Research Needs

Countering the UIT social engineering problem poses major challenges to organizations that must balance operational goals with security goals to maintain a competitive edge in the market. Because organizational policies and practices are resistant to change, it is a great challenge to keep up with the rapidly changing, increasingly sophisticated social engineering attacks.

These challenges suggest several research needs, including further study of organizational and human factors as well as additional case study data to increase the understanding of social engineering attacks. To advance the current practice and state of the art in computer and network defense, and especially safeguards against social engineering, the following research needs should be addressed.

7.2.1 Assessment of State of Practice and Effectiveness of Tools

Research should focus on development of cyber and network defense tools and antiphishing tools to better recognize and counter social engineering exploits. This research should not only focus on the availability of mitigation techniques but also on approaches to assessing the effectiveness of social engineering mitigation strategies. This research should include a survey of existing mitigation tools and approaches, identification of possible gaps in these defenses, and suggestions about technical or nontechnical means of filling these gaps.

Our efforts in the current project have brought us in touch with others in this field who are engaged in research as well as those who support client organizations through training programs and injection testing. The proposed research on effectiveness of mitigation strategies and tools (including training) may be performed in collaboration with other companies in this field, particularly those who have focused on the human element in cybersecurity. Collaborative research and a workshop with invited representatives from this community are possible approaches to carrying out this research.

7.2.2 Development of an Extensive UIT Database

A major roadblock to advancing our understanding and ability to counter UIT social engineering exploits is a dearth of accurate data from actual cases. By searching the internet, we have collected a small set of cases that contain limited details, but we expect that far more case information could be obtained directly from affected organizations. A self-reporting database is needed to collect and analyze incidents of social engineering. We

suggest that a feasibility analysis should be conducted to assess whether organizations could be motivated to self-report incidents, how the data may be collected anonymously and nonpunitively, and how the database can collect sensitive information from organizations such as government contractors and federally funded research and development centers (FFRDCs). The feasibility assessment also should study existing, successful databases such as the Defense Technical Information Center Independent Research and Development Collection, the Federal Aviation Administration's (FAA's) Aviation Safety Reporting System database, and the Department of Homeland Security (DHS) Cyberpilot Database. We also recommend conducting a needs assessment for such a database from the perspective of organizations and employees. The proposed information-sharing mechanism for social engineering incidents should be adopted across as wide a community as possible. The repository should track three types of information for every incident:

1. the actual malicious email itself (if that was the exploit vector)
2. data on the temporal ordering of pre- and post-exploit events
3. data on organizational factors and human factors that provide contextual information surrounding the incident

The proposed self-reporting incident database has numerous advantages:

- It would support trend analysis and provide targeted warning messages to other organizations that may be at risk. These warning messages may be formatted much like Food and Drug Administration (FDA) advisories so that these messages can be readily passed on to all company employees. Information in the warning messages may include type of attack vector, cues that the attack vector is a social engineering exploit, obfuscation techniques, appropriate responses to the attack vector, and contact information for further information. Analysis of trends in the collected data may include information used to target opportunities, people, and organizations; cues used in different attack vectors (e.g., emails, phone calls); obfuscation techniques; most common consequences of unmitigated attack (including attack patterns); and most common targets of the attacks.
- The dissemination of timely information about potential adversarial exploits may reduce the cost of exploits to potentially targeted organizations. Warning messages, following the FDA format (which has been empirically validated), are intended to be usable to the average non-expert in communicating information about socially engineered exploits. These messages are intended for rapid dissemination to all employees who communicate information about how to identify the threat and how to appropriately respond to it.
- To our knowledge, no database is collecting human factors information that can help researchers understand how adversaries leverage their knowledge of end users' psychological factors and cognitive limitations to craft their exploits. More detailed reporting of incident data would not only increase our understanding of how adversaries exploit human limitations but also document how people respond or do not respond to certain exploit cues.

To be sure, the proposed self-reporting incident database is not without challenges. An obstacle to success is reluctance of organizations to disclose information about exploits, especially organizations that do work in classified settings. Other similar databases exist that may successfully address this type of constraint (e.g., the Independent Research and Development (IR&D) databases at the Defense Technical Information Center (DTIC), the Defense Industrial Base (DIB) at The Department of Homeland Security, and possibly the Federal Bureau of Investigation's (FBI's) database for InfraGard organizations); however, we do not know the details on their incident collection, reporting, and dissemination. Other incident reporting databases, such as the FAA's Aviation Safety Reporting System, benefits from a federal mandate that all commercial airline pilots

report all aviation incidents, in exchange for asylum from punitive measures for reporting the incident. Federally mandating the reporting of cyber incidents is not in the purview of legislation, nor is it a likely possibility. Further discussion of the database and reporting concept is needed to work out details and overcome obstacles to their establishment. One possible approach to consider is a series of workshops held with stakeholders.

7.2.3 Detailed Analysis of UIT Incidents

Research is required to examine UIT incidents across a broad spectrum of participants in a comprehensive range of industries representing the full breadth of the economy. This research should focus on what factors are present in UIT incidents, how they have been handled by their respective organizations, and the motivation of those conducting UIT-HACK social engineering exploits. Our current effort was hampered in that we only had access to court transcripts and other third-party accounts of the incidents. Only through more detailed data collection, including collection of parameters that have been discussed in this report (but not represented in current reporting venues), will we be able to advance our understanding of UIT social engineering by applying the data analysis and conceptual modeling approaches described in this report.

Detailed analysis of UIT incidents—especially when informed by collected data that are relevant to possible behavioral and technical contributing factors—will help to identify more distinct conceptual patterns with which to characterize such incidents. If we can define more commonalities and distinct patterns across incidents, we will be better able to design more effective mitigation strategies. One challenge to pattern development generally is that the level of detail must be somewhat consistent across patterns to enable a valid comparative analysis. Not only is pattern identification and pattern reporting subjective, but the methodology employed for this type of work is typically not articulated in manuscripts, so we have little evidence for how patterns are being generated.

A way to address this problem is to use adapt a Contextual Design approach [Beyer 1998, Holtzblatt 2005] not only to guide data collection processes but also to provide a methodology to support data categorization and model building. The resulting models will ultimately describe different types of patterns (e.g., temporal event patterns, patterns of artifacts used). Contextual Design is a paradigm comprising qualitative data collection and data analysis methods to model the workflow of a system of people using technology. Contextual Design outputs designs or redesigns of systems built on explicit methodologies for data collection, analysis, and modeling. This paradigm follows the ethnographic, observational approach to research, which requires that the researchers divorce themselves from prior assumptions and allow the data to drive insights, hypotheses, and models. The step-by-step process of Contextual Design is well suited to addressing the level-of-detail pattern problem and methodology problem. Each step in the modeling process is standardized with naming conventions that include examples of the required level of detail. Contextual Design traditionally generates raw data through ethnographic interviews of individuals executing daily work tasks. The adaptation of the Contextual Design technique to the analysis of social engineering UIT incidents should include interviews of individuals involved in incidents; when interviews cannot be conducted, third-party sources must suffice.

The following outlines a high-level, step-by-step process for Contextual Design:

1. Collect data via interviews, court documents and transcripts, and other ethnographic methods.
2. Interpret the data.
3. Build cultural, physical (if possible), sequence, artifact, and workflow models.
4. Affinity diagram all data collected.
5. Walk the models and affinity.
6. List insights.
7. Conduct visioning (if the customer would like solutions to existing problems).

8 Recommendations

A challenge in conducting research on the contributing factors and mitigating strategies of socially engineered UIT incidents is the lack of peer-reviewed academic research on the topic. Additionally, the lack of quality reporting of socially engineered UIT incidents and case studies makes it difficult to study contributing factors; this is in part due to concerns about security, proprietary business practices, and litigation as well as the immaturity of reporting processes.

The use of deception and obfuscation in socially engineered UIT incidents presents special challenges for research aimed at developing effective mitigation strategies. For example, some phishing campaigns can be so well obfuscated that they appear 100% genuine to humans, and the adversarial success rate is very high. Other, less-obfuscated messages capitalize more on human limitations to succeed (e.g., highly fatigued employees may have lower performance thresholds). To add to the complexity, there is evidence that adversaries continually change their deceptive tactics. Regardless of a workforce's skill, savvy, or training, a phishing campaign always has a chance to succeed, especially because just one successful phishing campaign can penetrate a network. Nevertheless, the research community as well as responsible organizations and stakeholders have an obligation to continue research and information gathering to inform the development of effective training and mitigation tools.

Our review and analysis of research and case studies suggest the following strategies to reduce the effectiveness of social engineering attacks.

1. Continue to record demographic information as case studies are tabulated and entered into the UIT database. The records should include the demographic factors described in this report as well as the individual's role in the organization (e.g., function and position title).
2. Organizations should ensure that their management practices meet human factors standards that foster effective work environments to minimize stress (e.g., minimizing time pressure and optimizing workload) and encourage a healthy security culture. Because employees may perceive information security compliance as interfering with job functions, it is important for organizations to allocate a certain amount of employees' time to fulfilling the compliance requirements.
3. Organizations should develop and deploy effective training and awareness programs aimed at educating staff about social engineering scams, including learning objectives to help staff attend to phishing cues, identify deceptive practices, and recognize suspicious patterns of social engineering exploits. Training should also teach effective coping and incident management behaviors (ways to overcome one's own limitations and susceptibilities as well as appropriate responses to social engineering exploits).
4. The research and stakeholder community should develop mitigations that apply to specific attack phases as described in this report:
   - *Research and Open Source Intelligence phase*—Both the organization and individual employees may benefit from limiting online information that outsiders might exploit.
   - *Planning and Preparation phase*—Anticounterfeiting strategies that allow encrypted emails are well known but not commonly used.
   - *Launch Operation phase*—Improved employee training and awareness approaches should apply training that not only maintains staff vigilance and knowledge about the most current social

engineering tactics but also addresses human and organizational factors that may underlie vulnerabilities. Periodic injection testing and associated training should also be used.

- *Information Capture and Culmination and Exploitation phases*—Organizations should enable and maintain improved tools for computer and network defense cyber monitoring to keep up with the rapidly evolving kinds of exploits that adversaries use.

Countering the UIT social engineering problem poses major challenges to organizations that must balance operational goals with security goals to maintain a competitive edge in the market. Because organizational policies and practices are resistant to change, it is a great challenge to keep up with the rapidly changing, increasingly sophisticated social engineering attacks. Some social engineering campaigns may be so well crafted that they can defeat the organization's best countermeasures (e.g., training and policies). Attackers succeed even if only one employee succumbs to an exploit, so an organization's strategy to combat UIT social engineering must be comprehensive and include cybersecurity tools, security practices, and training.

Research is needed to further study possible contributing factors, particularly organizational and human factors. Additional case study data must be collected to increase understanding of characteristics of social engineering attacks. By characterizing and conceptually modeling the UIT social engineering problem, this report has sought to inform mitigation development efforts and identify research needs to more effectively combat UIT social engineering exploits.

Appendix A: Possible Contributing Factors in Social Engineering Susceptibility

Table 6 summarizes research findings related to social engineering susceptibility. Column 1 shows possible contributing factors, following the discussion in Section 4 of this report. Column 2 provides brief listings of the research findings that were cited in Section 4 (all citations are indicated as in Section 3 and listed in the References section). The research findings are identified as studies that are primarily experiments {E}, those that are primarily surveys {S}, those that are best characterized as theoretical or review papers {T}, or those that derive from news media or web-based articles {N}.

In addition, we have indicated, where possible, case studies that appear to have material that relates to these potential contributing factors. As noted in Section 5, case study data are sparse and typically do not contain direct mention of contributing factors, so in many cases we indicate these associations as inferred {I} versus documented {D}.

Table 6: Summary of Research Findings

Contributing Factors	Relevant Research Findings {E} = studies that fall mostly in experiments category {S} = studies that were primarily surveys {T} = theoretical or review papers {N} = news or web-based articles	Case Study Data by Case# {D} = documented {I} = inferred
Demographic Factors		
Age	• {E} Results of experiment are not conclusive in drawing a pattern for susceptibility to phishing [Mohebzada 2012]. • {E} No significant differences were found in phishing susceptibility between students, faculty, and staff in a university setting [Dhamija 2006]. • {E} 72% of 487 students aged 18–24 at Indiana University were successfully phished in an experiment [Jagatic 2007]. • {S} People aged 18–25 are more susceptible to phishing [Sheng 2010].	• {I} 1 • {I} 2 • {I} 24
Gender	• {E} Males and females were equally deceived in an initial phishing attack, and males were more likely to click on phishing links in a second attack and are more likely to provide personal information: results are not conclusive in drawing a pattern for susceptibility to phishing [Mohebzada 2012]. • {E} Women feel more comfortable with digital communication and may be more inclined to reply to emails with commercial offers or prizes [Halevi 2013]. • {E} Females are more susceptible to phishing than males [Halevi 2013]. • {S} Females are more susceptible to phishing than males; possible reason is that women have less technical experience than men [Sheng 2010].	• {D} 6 • {D} 9
Cultural	• {E} 8.74% of 10,917 students sampled from the American University of Sarjah in the United Arab Emirates fell for a phishing exploit [Mohebzada 2012]. • {E} Students were more susceptible to phishing attacks than faculty or staff, warning notices against phishing attempts were largely ignored, and users had difficulty recognizing the phishing schemes [Mohebzada 2012]. • {E} 7% of 200 students who participated in a study in Saudi Arabia responded to a phishing email [Alseadoon 2012]. • {E} Published results from a variety of studies report phishing response rate in the range between 3% and 11% in Western cultures, suggesting little, if any, cultural differences in phishing susceptibility [Dhamija 2006, Jakobsson 2006, Knight 2004, Mohebzada 2012].	

Contributing Factors	Relevant Research Findings {E} = studies that fall mostly in experiments category {S} = studies that were primarily surveys {T} = theoretical or review papers {N} = news or web-based articles	Case Study Data by Case# {D} = documented {I} = inferred
Personality traits	• {E} Neuroticism was more highly correlated to responding to a phishing email scheme [Halevi 2013]. • {E} Openness contributes to social engineering susceptibility [Alseadoon 2012, Halevi 2013]. • {E} People who are higher in normative commitment, more trusting, and more obedient to authority are more likely to succumb to social engineering [Workman 2008]. • {E} Low levels of conscientiousness predicted deviant workplace behavior such as breaking rules or behaving irresponsibly [Salgado 2002]. • {E} Responding to phishing emails represents an error in judgment that may be due to certain emotional biases; to examine this hypothesis, a study was performed that found personality traits associated with phishing, but only for females [Halevi 2013]. However, out of 100 participants, only 17 were female, so possible sampling errors may cast some doubt on the result. • {T} Extraversion may lead to increased phishing vulnerability [Parrish 2009]. • {T} Because openness is associated with technological experience and computer proficiency, people who score high on openness could be less susceptible to social engineering attacks [Parrish 2009]. • {T} Agreeableness may be the personality factor most often associated with social engineering susceptibility because of agreeable peoples' tendencies for trust, altruism, and compliance [Parrish 2009]. • {T} Higher levels of conscientiousness would result in individuals being more likely to follow training guidelines and less likely to break security policies [Parrish 2009].	• {D} 6 • {D} 14
Organizational factors		
Insufficient security systems, policies, and practices	• {E} Many people are not aware of phishing attacks [Mohebzada 2012]. • {S} Exposure to antiphishing education may have a large impact on phishing susceptibility [Sheng 2010]. • {E} Systems that are difficult to understand or to use are negatively perceived by users and are less likely to be used [Venkatesh 2003]. • {T} Easy-to-use passwords are not secure, but secure passwords are not easy to use [Zurko 1996]. • Security measures are often difficult and confusing for an average computer user, and errors caused by difficulty of security systems can yield serious consequences [Whitten 1999].	• {D} 5 • {D} 24 • {I} 1 • {D} 7 • {D} 28 • {D} 8 • {D} 9 • {D} 11 • {D} 20
Inadequate management and management systems	• {E} People are not well informed about the nature of the phishing threat or how to recognize social engineering schemes [Dhamija 2006, Mohebzada 2012]. • {S/E} Users who could correctly define phishing were less vulnerable to a phishing attack in a role-playing scenario, and participants who had experience with phishing websites were less likely to click on phishing links [Downs 2007]. • {T} Current antiphishing tools may not detect malicious websites that are well implemented [Erkkila 2011].	

Contributing Factors	Relevant Research Findings {E} = studies that fall mostly in experiments category {S} = studies that were primarily surveys {T} = theoretical or review papers {N} = news or web-based articles	Case Study Data by Case# {D} = documented {I} = inferred
Job pressure (time factors/deadline, task difficulty)	• {E} Users are more likely to respond to phishing emails in the presence of large email loads [Vishwanath 2011]. • {E} Time pressure negatively affects performance of even well-trained individuals [Lehner 1997]. • {E} Heavy and prolonged workload can cause fatigue, which adversely affects performance not only on simple tasks but also on more complex tasks [Soetens 1992]. • {E} Stressors in the workplace have a tendency to negatively impact human performance and increase errors, brought about through cognitive effects such as narrowing of attention (attending to fewer cues) [Houston 1969, Stokes 1994] and reduced working memory capacity [Davies 1982, Hockey 1986, Wachtel 1968].	
Human factors		
Lack of attention	• {E} Users do not pay attention to the source, grammar, and spelling used in a phishing email; instead they focus disproportionately on urgency cues [Vishwanath 2011]. • {E} Four cues in a phishing email could attract individual attention: subject line or title, email source, urgency cues, grammar and spelling [Vishwanath 2011]. • {E} Cues may be missed in the address bar and status bar [Dhamija 2006]. • {T} Users may not notice or read the security warnings or other security indicators and fail to notice the absence of security indicators (e.g., a padlock icon in the status bar) when they should be present [Erkkila 2011]. • {T} Even if antiphishing tools may technically work, users mostly ignore them, so they are inefficient [Erkkila 2011]. • {S/E} People tend to prefer cues in a site's content rather than more authoritative tools. Warnings and toolbars may use terms that are often not understood [Downs 2006]. • {E} 23% of 22 participants ignored browser-based security cues (address bar, status bar, SSL padlock icon), leading to incorrect choices 40% of the time [Dhamija 2006]. • {E} Visual deception practiced by phishers could fool even the most sophisticated users [Dhamija 2006]. • {E} A characteristic of successful phishing emails is a sense of urgency [Milletary 2005, Chandrasekaran 2006].	
Lack of knowledge/ memory failure	• {E} Users lack knowledge of the design inconsistencies that distinguish real and fake error messages [Sharek 2008]. • {E} Users have little awareness of potential risks involved in clicking fake popups [Sharek 2008]. • {T} Users lack basic understanding of the structure of the internet and computer systems in general [Erkkila 2011]. • {E}{S} Key factors are lack of knowledge about computer systems and security features (e.g., padlock icon) and lack of an understanding of URL/domain syntax and internet basics [Dhamija 2006, Downs 2006]. • {S/E} Users are less aware of social engineering attacks aimed at eliciting information directly from them [Downs 2006]. • {E} Key knowledge elements are knowledge about security features and understanding of URL/domain name syntax [Dhamija 2006].	• {I} 2 • {I} 1 • {I} 27 • {I} 9 • {I} 10

Contributing Factors	Relevant Research Findings {E} = studies that fall mostly in experiments category {S} = studies that were primarily surveys {T} = theoretical or review papers {N} = news or web-based articles	Case Study Data by Case# {D} = documented {I} = inferred
Faulty reasoning/ judgment	• {T} Users tend to ignore threats, thinking that they are an unlikely possibility [Sandouka 2009)]. • {T} Users underestimate the abilities of social engineers, and they generally hold misconceptions that organizational security systems are very secure by design [Sandouka 2009]. • {T} Users are encouraged by the salesperson mentality that the "client is always right," which allows social engineers to manipulate the friendly and helpful administrator or helpdesk [Sandouka 2009]. • {E} People's decisions tend to be biased and are not purely logical [Kahneman 1979]. • {E} Annoyance with popup messages may lead users to click on fake popups [Sharek 2008]. • {T} Users may think that they do not need redundant security features that slow down their job and that security risks in the internet are over-hyped [Erkkila 2011]. • {T} Users who rely on habituation to process cues about phishing in email messages are not processing messages cognitively at sufficient depth to detect some fairly obvious cues [Watters 2009]. • {E} Users consider emails to be more phishy than webpages, and webpages to be more phishy than phone calls [Jakobsson 2007].	• {I} 6
Risk tolerance/risk perception	• {E} Warning notices against phishing attempts sent to users were largely ignored [Mohebzada 2012]. • {E} People who are more engaged with Facebook activity have less restrictive privacy settings [Halevi 2013]. • {S} The more risk-averse a participant is, the less likely he or she will fall for phishing [Sheng 2010].	
Casual values/attitudes about compliance	• {S} Employee attitudes, normative beliefs, and habits are major determinants of intention to comply with information-system security policy [Pahnila 2007, Bulgurcu 2010]. • {S} Sanctions did not significantly influence employees' intention to comply, and awards did not have a significant effect on actual compliance [Pahnila 2007]. • {S} Beliefs about overall assessment of consequences are immediate antecedents of attitude; thus, factors that motivate employees to comply with the information-system security policies extend beyond sanctions and rewards [Bulgurcu 2010]. • {S} The impact of the cost of compliance is as strong as the impacts of the benefit of compliance and the cost of noncompliance [Bulgurcu 2010]. • {S} Creating a security-aware culture within the organization will improve information security [Bulgurcu 2010].	
Stress/anxiety	• {E} Time pressure negatively affects performance of even well-trained individuals [Lehner 1997]. • {T} Job stress negatively affects employee performance [Leka 2004].	
Impaired physical status	• {E} A study of neurocognitive impairment reported impaired neurocognitive performance in approximately two-thirds of patients who entered a14-day inpatient substance abuse unit; the most frequently compromised areas of functioning involved attention, memory, calculation, abstraction, ability to follow complex commands, and visuospatial skills [Meek 1989]. • {P} Abuse of drugs and alcohol may be associated with loss of productivity, among other problems [HealthyPeople.gov 2013]. • {E} Dopamine is more pervasive in the brains of risk-takers, or they have fewer dopamine-inhibiting receptors [Zald 2008]. • {E} Dopamine plays a role in the amount of risks that people take [Park 2008].	

Appendix B: Case Study Material

To better understand the scope and variety of social engineering exploits, we used a case study approach that collects, abstracts, and reports on actual incidents. Using a set of descriptive parameters, incidents are summarized succinctly and expressed in a clear and consistent manner for informal review. These parameters, borrowed from the Phase 1 research, are defined as follows:

- INCIDENT ID: <ID #>
- INDUSTRY: <classification of organization>
- STAGING: <single, multiple>
- INCIDENT: <description of the how social engineering was used, across multiple stages where applicable>
- BREACH: <type of loss or compromise>
- OUTCOME: <organizational status resulting from breach>
- RESPONSE: <specific action taken in response to the breach>
- REFERENCES: <URLs, or references to sources of incident descriptions>

These parameters are used to classify all social engineering cases. Case study data collected to date have been entered into a UIT database. To preserve organizational privacy and anonymity, the information reported in this appendix does not include names or other identifiable information, including website URLs and citations of news articles or legal judgments. As such, this report omits the REFERENCES field of the incident summaries. Also, because some descriptive information was deleted or modified to preserve privacy, associations with possible contributing factors may be less apparent.

In all, there are 28 cases in our UIT social engineering database. All the cases were found online, such as through search engines. Three of the cases (10.7%) have more than one source reference. A breakdown of the sources is as follows:

- news articles: 25/28 (89.3%)
- journal publications: 1/28 (3.6%)
- blog: 1/28 (3.6%)
- other: 1/28 (3.6%)

INCIDENT ID:	1
INDUSTRY:	Manufacturing
STAGING:	Single
INCIDENT:	Malware to attack the victim organization and other companies was spread through a website for software developers. The website advertised a java plug-in that could be installed on desktops.
BREACH:	A number of employees installed the malware disguised as a java plug-in. The systems impacted by the attack were low in number and isolated from the network.
OUTCOME:	The victim organization worked with law enforcement to find the source of the malware and released a tool to remove Java malware.
RESPONSE:	The victim organization's native anti-malware software was updated to automatically detect and isolate the malware.

INCIDENT ID:	2
INDUSTRY:	Banking and Finance
STAGING:	Single
INCIDENT:	Fake emails containing malware were being sent to employees of a financial institution. Six employees opened the fake email and downloaded malware. The malware was not detected by virus protection software on the employees' computers.
BREACH:	Confidential information was disclosed. The malware did not spread to other parts of the network because the six employees did not have administrative access.
OUTCOME:	Opening the fake emails was a violation of company's information security policy.
RESPONSE:	Unknown.

INCIDENT ID:	3
INDUSTRY:	Information Technology
STAGING:	Single
INCIDENT:	A number of employees at the victim organization fell victim to a spear phishing attack. The phishing attack included emails from various government and commercial organizations to the employees. Though the email was sent to a large number of employees, only about 1% of those employees executed the malware.
BREACH:	The emails executed malware that exfiltrated megabytes of data.
OUTCOME:	Unknown.
RESPONSE:	Unknown.

INCIDENT ID:	4
INDUSTRY:	Defense Industrial Base
STAGING:	Single
INCIDENT:	A fake voice over internet protocol (VoIP) client was downloaded by the victim organization. The victim organization downloaded the client under the belief that the client provided encrypted communication. The software was a remote administration tool that allowed the attackers to turn on the infected computer's webcam and remotely monitor activity. The software could also be used to record keystrokes and steal passwords.
BREACH:	The victim organization downloaded a fake VoIP client that gave the attackers remote access to machines in that organization.
OUTCOME:	The attackers had full access to the victim organization's computing systems.
RESPONSE:	Unknown.

INCIDENT ID:	5
INDUSTRY:	Banking and Finance
STAGING:	Multiple
INCIDENT:	The attacker sent a phishing email impersonating the victim organization's bank. The email requested information to address security concerns. The employee at the victim organization went to the webpage and entered confidential information.
BREACH:	The attack resulted in the disclosure of credentials and passwords that enabled outsiders to transfer funds to accounts in several countries.
OUTCOME:	The bank was able to recover approximately 70% of what the victim organization lost.
RESPONSE:	The victim organization recovered the remainder of lost funds in a court settlement resulting from a lawsuit filed against bank.

INCIDENT ID:	6
INDUSTRY:	Government
STAGING:	Multiple
INCIDENT:	An employee at the victim organization browsed a website unrelated to work and inadvertently downloaded malware. The malware ran a key logger on the employee's computer. The malware went undetected for a period of five months and was discovered when the employee was terminated and the employee's hard drive was scanned.
BREACH:	The attackers gained access to over 2,000 people's personally identifiable information (PII).
OUTCOME:	The victim organization notified those impacted by the data breach and offered credit monitoring services.
RESPONSE:	The department's employees are now prohibited from doing any personal web surfing at work.

INCIDENT ID:	7
INDUSTRY:	Government
STAGING:	Multiple
INCIDENT:	Attackers sent phishing emails to employees at the victim organization. At least one employee downloaded malware from the email. The malware copied and exfiltrated the employee's username and password. The attackers then used the employee's credentials to access information systems and exfiltrate data from the network.
BREACH:	The attackers used the employee's credentials to access the PII of over three million people.
OUTCOME:	The victim organization brought in a third party for review of the incident and changed login practices.
RESPONSE:	Organization changed their login practices.

INCIDENT ID:	8
INDUSTRY:	Financial Services
STAGING:	Multiple
INCIDENT:	An employee at the victim organization replied to a phishing email. The employee believed the email to be from a financial services provider. The employee downloaded and installed keystroke-logging malware. The malware captured the employee's credentials.
BREACH:	The attackers were able to transfer hundreds of thousands of dollars using the employee's credentials.
OUTCOME:	The victim organization filed a lawsuit claiming the financial institution did not follow correct security practices. The financial institution countersued claiming the victim organization had declined additional security measures.
RESPONSE:	Common recommendation: Victim organization should adopt dual authorization: Two designated individuals must authorize any transfer.

INCIDENT ID:	9
INDUSTRY:	Education
STAGING:	Single
INCIDENT:	Attackers at the victim organization created a false request for credentials. The attackers gained access when the false prompt was presented to an employee with administrative rights. The attackers then viewed confidential information located at the victim organization.
BREACH:	The attackers used the employee's credentials to view confidential information.
OUTCOME:	The attackers did not continue their attack after they initially gained access.
RESPONSE:	Training was provided to members of the victim organization.

INCIDENT ID:	10
INDUSTRY:	Government
STAGING:	Multiple
INCIDENT:	An employee at the victim organization received a phishing email and entered data into a fraudulent website.
BREACH:	The attacker used the employee's information to gain access to other employees' PII.
OUTCOME:	The data breach was contained to employee information and did not compromise any trading or market data. Law enforcement was contacted, and additional security controls were imposed.
RESPONSE:	Staff will receive increased training, especially those who handle PII. Employees will receive identity protection from a credit-monitoring company.

INCIDENT ID:	11
INDUSTRY:	Education
STAGING:	Multiple
INCIDENT:	Seven employees at the victim organization entered credentials into a fraudulent site after receiving a phishing email.
BREACH:	The breach compromised the security of employee email accounts. The affected emails contained PII and financial information about as many as 500 individuals.
OUTCOME:	The compromised accounts were restored from backup sources.
RESPONSE:	The institution informed government officials and retained computer forensic and breach notification experts. Staff was retrained at the organization.

INCIDENT ID:	12
INDUSTRY:	Health Care
STAGING:	Single
INCIDENT:	A subcontractor for the victim organization responded to a phishing scam that allowed remote access to the victim organization's customer data.
BREACH:	More than 1,000 medical records were compromised as well as additional PII.
OUTCOME:	Customers were notified of the potential breach.
RESPONSE:	Free credit monitoring was provided for customers.

INCIDENT ID:	13
INDUSTRY:	Information Technology
STAGING:	Multiple
INCIDENT:	An employee of the victim organization fell victim to a targeted phishing attack. The employee downloaded malware that recorded and sent the employee's password to the attackers. The attackers were then able to use the employee's access to copy the organization's customer list.
BREACH:	The attackers used the victim organization's customer list to send targeted phishing emails to the victim organization's customers.
OUTCOME:	Some of the victim organization's passwords were obtained by attackers.
RESPONSE:	Unknown.

INCIDENT ID:	14
INDUSTRY:	Government
STAGING:	Multiple
INCIDENT:	The attackers compromised the accounts of top officials in the victim organization. The attackers then used these accounts to send requests to different directors at the organization requesting passwords for information systems. The attackers also used the compromised accounts to send malware to other members of the staff. The malware searched for sensitive information and sent it back to the attackers over the internet.
BREACH:	The attackers compromised the accounts of high-level employees at the victim organization.
OUTCOME:	The victim organization shut down internet connectivity to prevent data exfiltration.
RESPONSE:	Unknown.

INCIDENT ID:	15
INDUSTRY:	News
STAGING:	Single
INCIDENT:	The attackers sent a phishing email to employees at a victim organization. The employees entered credentials to a social media site. The attackers used the employee's social media to report fake news.
BREACH:	The attackers were able to use a phishing attack to gain the credentials to an employee's social media account.
OUTCOME:	The fake news caused a panic in financial markets. Once the news was found out to be fake, the market stabilized.
RESPONSE:	Unknown.

INCIDENT ID:	16
INDUSTRY:	Government
STAGING:	Multiple
INCIDENT:	Attackers sent fake emails to employees at the victim organization. The emails contained a trojan and malware that stole passwords and exfiltrated data from the employee's computers.
BREACH:	Employees at the victim organization downloaded malware that exfiltrated data.
OUTCOME:	Over two gigabytes of data was exfiltrated from the victim organization.
RESPONSE:	Unknown.

INCIDENT ID:	17
INDUSTRY:	Education
STAGING:	Multiple
INCIDENT:	The attackers used a phishing email to collect credentials of employees at the victim organization. The attackers then used the credentials to access the employees' accounts at a credit union.
BREACH:	Attackers used a phishing email to collect the employee credit union credentials at the victim organization.
OUTCOME:	Over 50 credentials were disclosed.
RESPONSE:	Unknown.

INCIDENT ID:	18
INDUSTRY:	News
STAGING:	Multiple
INCIDENT:	The attackers sent a phishing email to employees at the victim organization. The email provided a link that gave employees' credentials to attackers. Once the attackers gained access to one employee's email account, the attackers used the account to send more phishing emails to other members of the organization. After receiving a malicious email from the originally compromised employee's account, a second employee fell for the phishing exploit and provided credentials to the attackers. The attackers used the second employee's credentials to access the victim organization's accounts on multiple social media sites.
BREACH:	The attackers compromised multiple accounts at the victim organization and sent multiple phishing emails to employees within the organization.
OUTCOME:	The victim organization's social media accounts were hacked, and multiple employees were victims of a phishing attack.
RESPONSE:	Unknown.

INCIDENT ID:	19
INDUSTRY:	Logistics
STAGING:	Single
INCIDENT:	Attackers sent a phishing email to an employee at the victim organization. The employee acted on the phishing email resulting, in a compromised machine in the network. The employee's computer contained PII of other employees in the organization.
BREACH:	One computer at the victim organization was compromised. It is unknown whether the attacker exfiltrated data from the machine.
OUTCOME:	The victim organization took the computer offline and performed an investigation to ensure that no other computers were impacted.
RESPONSE:	The victim organization offered identity theft protection for those whose PII was on the compromised machine.

INCIDENT ID:	20
INDUSTRY:	Health Care
STAGING:	Multiple
INCIDENT:	Attackers sent employees at the victim organization an email that appeared to come from a trusted source. The email contained a link that requested the employee's credentials. Multiple employees fell victim to the attack, and the breach was detected on the same day it occurred.
BREACH:	Attackers accessed records containing PII of more than 2,000 customers.
OUTCOME:	The victim organization notified customers whose data was accessed by hackers.
RESPONSE:	The victim organization performed an investigation and provided training to prevent another attack from occurring in the future.

INCIDENT ID:	21
INDUSTRY:	Information Technology
STAGING:	Multiple
INCIDENT:	Attackers compromised the victim organization, a website-hosting company, with a phishing attack.
BREACH:	There was no evidence that customer data was stolen.
OUTCOME:	Unknown
RESPONSE:	The victim organization contacted customers and provided advice to limit the impact in case customer data had been stolen.

INCIDENT ID:	22
INDUSTRY:	Information Technology
STAGING:	Single
INCIDENT:	An employee at the victim organization downloaded an attachment with a zero-day exploit that led to the installation and execution of malware. Employees at the victim organization discovered the attack was an attempt to steal one of the organization's new products.
BREACH:	Installation of malware allowed the attackers to exfiltrate data.
OUTCOME:	Data was exfiltrated from the victim organization.
RESPONSE:	The vulnerability was patched, and the worker was not fired. The company has since increased the amount of security surrounding their network, and it is actively working to identify zero-day exploits.

INCIDENT ID:	23
INDUSTRY:	Social Engineering
STAGING:	Multiple
INCIDENT:	Attackers used social engineering techniques to convince a domain registrar to change the default email account associated with a financial institution. The attackers also convinced the domain registrar to reset the default password.
BREACH:	After the attackers gained access to the financial institution's web servers, the attackers denied service to the financial institution's employees and stole over $10,000.
OUTCOME:	The attack shut down the financial organization for a week.
RESPONSE:	The financial institution implemented two-factor authentication.

INCIDENT ID:	24
INDUSTRY:	Government – Federal
STAGING:	Multiple
INCIDENT:	Attackers sent a phishing email regarding HR benefits to employees at the victim organization. The phishing email downloaded malicious code that exploited a zero-day vulnerability. The malware was designed to compromise a system and exfiltrate data. The malware was programmed to erase itself if it failed to compromise a system.
BREACH:	Limited data was exfiltrated from the organization, but the victim organization failed to initially recognize dormant malicious code. This incident was classified as an advanced persistent threat (APT) due to the nature of the breach.
OUTCOME:	This was the second successful phishing attack at the victim organization. Learning lessons from the earlier incident, the victim organization disconnected internet access after network administrators discovered that data was being externally siphoned from a server. After initial shutdown, the victim organization allowed external email but blocked attachments.
RESPONSE:	Though there had been extensive training after the first attack, the organization's long-term response to the second attack is unknown.

INCIDENT ID:	25
INDUSTRY:	IT domain registration
STAGING:	Multiple
INCIDENT:	Users of a website-hosting service began receiving phishing emails. The phishing attack used rogue messages masquerading as alerts about account load limits being exceeded. The attack directed recipients to a rogue website where PII was captured.
BREACH:	It is unknown if hosting website data was compromised, but users may have disclosed credit card information, delivery addresses, and phone numbers on the rogue site.
OUTCOME:	An unknown number of customers were compromised.
RESPONSE:	The hosting firm alerted users that it was receiving a high number of phishing emails.

INCIDENT ID:	26
INDUSTRY:	Financial
STAGING:	Multiple
INCIDENT:	A phishing attack compromised thousands of customers and stole over $30,000,000. The attackers used advanced phishing techniques that negatively impacted business partnerships.
BREACH:	Unknown.
OUTCOME:	Unknown.
RESPONSE:	Unknown.

INCIDENT ID:	27
INDUSTRY:	Telecommunications
STAGING:	Single
INCIDENT:	Attackers sent emails to customers that are designed to look like legitimate communications from businesses.
BREACH:	Once users click on an embedded link, they are redirected to compromised websites that point to the malicious sites hosting the exploit code.
OUTCOME:	The malware is downloaded to the customers' systems and can transmit key strokes or other information to the attackers.
RESPONSE:	Generally, companies have pages on their websites that highlight recent scams. Some organizations will email their customers to alert them of scams.

INCIDENT ID:	28
INDUSTRY:	Financial Services
STAGING:	Single
INCIDENT:	Attackers sent phishing emails to the customers of a payment processing company. During the incident, a number of its customers received an email warning them that they needed to download a Web browser plug-in in order to maintain uninterrupted access to a website. The plug-in was instead malicious software designed to steal the victim's usernames and passwords. The attackers targeted the customers by name in the body of the message. The phishing message also reference the recipient's username and a portion of his or her password for the site. The attackers had obtained the customer data to craft the phishing attacks through a direct attack on the company's servers. A second attack occurred two weeks later.
BREACH:	The payment processor did not know the total number of customer accounts stolen or how many customers had provided login information following the phishing attack.
OUTCOME:	The payment processor hired independent computer forensic experts and worked with federal law enforcement investigators.
RESPONSE:	The payment processor temporarily shut down its site and instituted new security measures to protect client information, such as requiring users to change their passwords.

References

URLs are valid as of the publication date of this document.

[AlgoSec 2013]
AlgoSec. *The State of Network Security 2013: Attitudes and Opinions*. AlgoSec, Inc., 2013.

[Alseadoon 2012]
Alseadoon, I.; Chan, T.; Foo, E.; & Nieto, J. G. "Who Is More Susceptible to Phishing Emails?: A Saudi Arabian Study." *Proceedings of the 23rd Australasian Conference on Information Systems*. Geelong, Australia, Dec. 3-5, 2012. ACIS, 2012. http://dro.deakin.edu.au/view/DU:30049075

[APWG 2005]
Anti-Phishing Working Group (APWG). *Phishing Trends Report, December 2005*. APWG, 2005. http://www.antiphishing.org/reports/apwg_report_DEC2005_FINAL.pdf

[Beyer 1998]
Beyer, H. & Holtzblatt, K. *Contextual Design: Defining Customer-Centered Systems*. Academic Press, 1998.

[Bijou 2013]
Bijou, R. Examining the Cyber Kill Chain. http://www.rbijou.com/2013/03/15/examining-the-cyber-kill-chain/ (2013).

[Bulgurcu 2010]
Bulgurcu, B.; Cavusoglu, H.; & Benbasat, I. "Information Security Policy Compliance: An Empirical Study of Rationality-Based Beliefs and Information Security Awareness." *MIS Quarterly 34*, 3 (Sep. 2010): 523-548.

[Cacioppo 1986]
Cacioppo, J. T.; Petty, R. E.; Kao, C. F.; & Rodriguez, R. "Central and Peripheral Routes to Persuasion: An Individual Difference Perspective." *Journal of Personality and Social Psychology 51*, 5 (Nov. 1986): 1032-1043.

[CERT 2013]
CERT Insider Threat Team. *Unintentional Insider Threats: A Foundational Study* (CMU/SEI-2013-TN-022). Software Engineering Institute, Carnegie Mellon University, May 2013. http://resources.sei.cmu.edu/library/asset-view.cfm?assetID=58744

[Chandrasekaran 2006]
Chandrasekaran, M.; Narayanan, K.; & Upadhyaya, S. "Phishing Email Detection Based on Structural Properties," 1-7. *Proceedings of the 9th Annual NYS Cyber Security Conference*. Albany, NY, June 2006.

[Cialdini 2001]
Cialdini, R. B. *Influence: Science and Practice*. Allyn and Bacon, 2001.

[Cloppert 2009]
Cloppert, M. "Security Intelligence: Attacking the Cyber Kill Chain." SANS Computer Forensics, October 14, 2009. http://computer-forensics.sans.org/blog/2009/10/14/security-intelligence-attacking-the-kill-chain

[Davies 1982]
Davies, D. R. & Parasuraman, R. *The Psychology of Vigilance*. Academic Press, 1982.

[Dekker 2002]
Decker, S. *The Field Guide to Human Error Investigations*. Ashgate, 2002.

[Dhamija 2006]
Dhamija, R.; Tygar, J. D.; & Hearst, M. "Why Phishing Works," 581-590. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '06*. New York, NY, Apr. 2006. ACM, 2006. http://dl.acm.org/citation.cfm?id=1124861

[Digman 1990]
Digman, J. M. "Personality Structure: Emergence of the Five-Factor Model." *Annual Review of Psychology 41* (1990): 417-440.

[Downs 2006]
Downs, J. S.; Holbrook, M. B.; & Cranor, L. F. "Decision Strategies and Susceptibility to Phishing," 79-90. *Proceedings of the Second Symposium on Usable Privacy and Security*. Pittsburgh, PA, July 2006. ACM, 2006.

[Downs 2007]
Downs, J. S.; Holbrook, M. B.; & Cranor, L. F. "Behavioral Response to Phishing Risk (Institute for Software Research, Paper 35)." *APWG eCrime Researchers Summit*. Pittsburgh, PA, Oct. 2007.

[Erkkila 2011]
Erkkila, J. "Why We Fall for Phishing." *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI 2011*. Vancouver, BC, Canada, May 2011. ACM, 2011.

[Goldberg 1971]
Goldberg, L. R. Ch. XIV, "A Historical Survey of Personality Scales and Inventories," 293-336. *Advances in Psychological Assessment* (Vol. 2), Science and Behavior Books, 1971.

[Goldberg 2006]
Goldberg, L. R.; Johnson, J. A.; Eber, H. W.; Hogan, R.; Ashton, M. C.; Cloninger, C. R.; & Gough, H. G. "The International Personality Item Pool and the Future of Public-Domain Personality Measures." *Journal of Research in Personality 40*, 1 (Feb. 2006): 84-96.

[Granger 2001]
Granger, S. "Social Engineering Fundamentals, Part I: Hacker Tactics." *SecurityFocus*, Dec. 17, 2001. http://www.securityfocus.com/infocus/1527

[Halevi 2013]
Halevi, T.; Lewis, J.; & Memon, N. *Phishing, Personality Traits and Facebook*. Cornell University Library, 2013. http://arxiv.org/abs/1301.7643

[HealthyPeople.gov 2013]
HealthyPeople.gov. *Substance Abuse*. http://healthypeople.gov/2020/LHI/substanceAbuse.aspx (2013).

[Hockey 1986]
Hockey, G. R. J. "Changes in Operator Efficiency as a Function of Environmental Stress, Fatigue, and Circadian Rhythms." *Handbook of Perception and Human Performance (Vol. II)*. Wiley, 1986.

[Holtzblatt 2005]
Holtzblatt, K.; Wendell, J. B.; & Wood, S. *Rapid Contextual Design*: *A How-To Guide to Key Techniques for User-Centered Design*. Elsevier, 2005.

[Houston 1969]
Houston, B. K. "Noise, Task Difficulty, and Stroop Color-Word Performance." *Journal of Experimental Psychology 82*, 2 (1969): 403-404.

[Hutchins 2011]
Hutchins, E. M.; Cloppert, M. J.; & Amin, R. M. "Intelligence-Driven Computer Network Defense Informed by Analysis of Adversary Campaigns and Intrusion Kill Chains." *6th Annual International Conference on Information Warfare and Security*. Washington, DC, Mar. 2011. Academic Publishing International, 2011.
http://www.lockheedmartin.com/content/dam/lockheed/data/corporate/documents/LM-White-Paper-Intel-Driven-Defense.pdf

[Jagatic 2007]
Jagatic, T. N.; Johnson, N. A.; Jakobsson, M.; & Menczer, F. "Social Phishing." *Communications of the ACM 50*, 10 (Oct. 2007): 94-100.

[Jakobbson 2005]
Jakobbson, M. "Modeling and Preventing Phishing Attacks," 89. *Financial Cryptography and Data Security: 9th International Conference (FC 2005), Revised Papers.* Roseau, Dominican Republic, Feb. 2005. Springer Berlin Heidelberg, 2005.
http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.64.1926&rep=rep1&type=pdf

[Jakobsson 2006]
Jakobsson, M. & Ratkiewicz, J. "Designing Ethical Phishing Experiments: A Study of (ROT13) Ronl Query Features," 513-522. *Proceedings of the 15th International Conference on World Wide Web*. Edinburgh, Scotland, May 2006. ACM, 2006.

[Jakobsson 2007]
Jakobsson, M. "The Human Factor in Phishing." *Privacy & Security of Consumer Information'07.* http://markus-jakobsson.com/papers/jakobsson-psci07.pdf

[John 1999]
John, O. P. & Srivastava, S. "The Big Five Trait Taxonomy: History, Measurement, and Theoretical Perspectives," 102-138. *Handbook of Personality: Theory and Research (2nd edition)*. Guilford, 1999.

[Kahneman 1979]
Kahneman, D. & Tversky, A. "Prospect Theory: An Analysis of Decision under Risk." *Econometrica 47,* 2 (Mar. 1979): 263-292. http://www.jstor.org/discover/10.2307/1914185?uid=3739960&uid=2&uid=4&uid=3739256&sid=21102164801913

[Knight 2004]
Knight, W. "Goin' Phishing?" *Infosecurity Today 1,* 4 (July-Aug. 2004): 36-38. http://www.sciencedirect.com/science/article/pii/S1742684704000898

[Laribee 2006a]
Laribee, L. "Development of Methodical Social Engineering Taxonomy." Master's thesis, Naval Postgraduate School, 2006.

[Laribee 2006b]
Laribee, L.; Barnes, D. S.; Rowe, N. C.; & Martell, C. H. "Analysis and Defensive Tools for Social-Engineering Attacks on Computer Systems," 388-389. *Information Assurance Workshop (IAW), 2006 IEEE*. West Point, NY, June 2006. IEEE, 2006. http://ieeexplore.ieee.org/xpl/login.jsp?tp=&arnumber=1652125&url=http%3A%2F%2Fieeexplore.ieee.org%2Fxpls%2Fabs_all.jsp%3Farnumber%3D1652125

[Lehner 1997]
Lehner, P.; Seyed-Solorforough, M.; O'Connor, M. F.; Sak, S.; & Mullin, T. "Cognitive Biases and Time Stress in Team Decision Making." *IEEE Transactions on Systems, Man, and Cybernetics Part A: Systems & Humans 27,* 5 (Sep.1997): 698-703.

[Leka 2004]
Leka, S.; Griffiths, A.; & Cox, T. *Work Organization and Stress: Systematic Problem Approaches for Employers, Managers, and Trade Union Representatives.* Protecting Workers Health Series, No. 3. World Health Organization, 2004. http://www.who.int/occupational_health/publications/pwh3rev.pdf

[McCombie 2010]
McCombie, S. & Pieprzyk, J. "Winning the Phishing War: A Strategy for Australia," 79-86. *2010 Second Cybercrime and Trustworthy Computing Workshop (CTC)*. Ballarat, Victoria, Australia, July 2010. Conference Publishing Services, 2010. http://ieeexplore.ieee.org/xpl/login.jsp?tp=&arnumber=5615076&url=http%3A%2F%2Fieeexplore.ieee.org%2Fiel5%2F5613989%2F5614937%2F05615076.pdf%3Farnumber%3D5615076

[McCrae 1992]
McCrae, R. R. & John, O. P. "An Introduction to the Five-Factor Model and Its Applications." *Journal of Personality 60,* 2 (June 1992): 175-215.

[Meadows 2008]
Meadows, D. *Thinking in Systems: A Primer*. Chelsea Green Publishing, 2008.

[Meek 1989]
Meek, P. S.; Clark, H. W.; & Solana, V. L. "Neurocognitive Impairment: The Unrecognized Component of Dual Diagnosis in Substance Abuse Treatment." *Journal of Psychoactive Drugs 21,* 2 (Apr.-June 1989): 153-160.

[Miller 2005]
Miller, K. *Communication Theories: Perspectives, Processes, and Contexts.* McGraw-Hill, 2005.

[Milletary 2005]
Milletary, J. *Technical Trends in Phishing Attacks*. US-CERT, Department of Homeland Security, 2005. http://www.us-cert.gov/security-publications/technical-trends-phishing-attacks

[Mitnick 2002]
Mitnick, K. & Simon, W. L. *The Art of Deception: Controlling the Human Element of Security.* Wiley, 2002.

[Mohebzada 2012]
Mohebzada, J. G.; El Zarka, A.; Bhojani, A. H.; & Darwish, A. "Phishing in a University Community," 249-254. *2012 International Conference on Innovations in Information Technology (IIT)*. Abu Dhabi, United Arab Emirates, Mar. 2012. IEEE, 2012.

[Myers 2013]
Myers, L. *Cyber Kill Chain Is a Great Idea, But Is It Something Your Company Can Implement?* Infosec Institute, 2013. http://resources.infosecinstitute.com/cyber-kill-chain-is-a-great-idea-but-is-it-something-your-company-can-implement/

[NIST 2002]
National Institute of Standards and Technology (NIST). *Risk Management Guide for Information Technology Systems (Special Publication 800-30)*. U.S. Department of Commerce, 2002.

[O'Brien 2005]
O'Brien, T. L. "Gone Spear-Phishin'." *The New York Times.* December 4, 2005. http://www.nytimes.com/2005/12/04/business/yourmoney/04spear.html?pagewanted=1&ei=5088&en=2f313fc4b55b47bf&ex=1291352400&partner=rssnyt&emc=rss&_r=0

[Pahnila 2007]
Pahnila, S.; Siponen, M.; & Mahmood, A. "Employees' Behavior Towards IS Security Policy Compliance," 156b. *40th Annual Hawaii International Conference on System Sciences*, *HICSS 2007*. Waikoloa, Big Island, HW, Jan. 2007. IEEE, 2007.

[Park 2008]
Park, A. "Why We Take Risks—It's the Dopamine." *Time*, Dec. 30, 2008. http://www.time.com/time/health/article/0,8599,1869106,00.html

[Parrish 2009]
Parrish, Jr., J. L.; Bailey, J. L.; & Courtney, J. F. "A Personality Based Model for Determining Susceptibility to Phishing Attacks," 285-296. *2009 Southwest Decision Sciences Institute*, Feb. 2009. SWDSI, Oklahoma City, Oklahoma, 2009.

[Paunonen 2001]
Paunonen, S. V. & Ashton, M. C. "Big Five Factors and Facets and the Prediction of Behavior." *Journal of Personality and Social Psychology 81,* 3 (Sep. 2001): 524-539.

[Peltier 2006]
Peltier, T. R. "Social Engineering: Concepts and Solutions." *Information Systems Security 15,* 5 (Nov. 2006): 13-21.

[Pond 2003]
Pond, D. J. & Leifheit, K. R. "End of an Error." *Security Management 47,* 5 (2003): 113-117.

[Salgado 2002]
Salgado, J. "The Big Five Personality Dimensions and Counterproductive Behaviors." *International Journal of Selection and Assessment 10,* 1-2 (Mar. 2002): 117-125.

[Sandouka 2009]
Sandouka, H.; Cullen, A. J.; & Mann, I. "Social Engineering Detection Using Neural Networks," 273-278. *International Conference on CyberWorlds 2009 (CW'09)*. Bradford, Yorkshire, United Kingdom, Sep. 2009. IEEE, 2009. http://ieeexplore.ieee.org/xpl/freeabs_all.jsp?arnumber=5279574

[Sastry 1989]
Sastry, A. "Archetypal Self-Reinforcing Structures in Organizations: A System Dynamics Perspective of Cognitive, Social, and Institutional Processes." *Conference Proceedings: The 7th International Conference of the System Dynamics Society.* Stuttgart, Germany, June 1989.

[Sharek 2008]
Sharek, D.; Swofford, C.; & Wogalter, M. "Failure to Recognize Fake Internet Popup Warning Messages," 557-560. *Proceedings of the Human Factors and Ergonomics Society 52nd Annual Meeting.* New York, NY, Sep. 2008. SAGE Publications, 2008.

[Sheng 2007]
Sheng, S.; Magnien, B.; Kumaraguru, P.; Acquisti, A.; Cranor, L.; Hong, J.; & Nunge, E. "Anti-Phishing Phil: The Design and Evaluation of a Game That Teaches People Not to Fall for Phish," 88-99. *Proceedings of the Third Symposium on Usable Privacy and Security (SOUPS).* Pittsburgh, PA, July 2007. ACM, 2007.

[Sheng 2010]
Sheng, S.; Holbrook, M.; Kumaraguru, P.; Cranor, L.; & Downs, J. "Who Falls for Phish? A Demographic Analysis of Phishing Susceptibility and Effectiveness of Interventions," 373-382. *28th ACM Conference on Human Factors in Computing Systems*. Atlanta, GA, April 2010. Interaction Design Foundation, 2010.

[Soetens 1992]
Soetens, E.; Hueting, J.; & Wauters, F. "Traces of Fatigue in an Attention Task." *Bulletin of the Psychonomic Society 30,* 2 (Aug. 1992): 97-100.

[Staddon 1996]
Staddon, J. & Higa, J. "Multiple Time Scales in Simple Habituation." *Psychological Review 103,* 4 (Oct. 1996): 720-733.

[Staw 1989]
Staw, B. M. & Ross, J. "Understanding Behavior in Escalation Situations." *Science 246,* 4927 (Oct. 1989): 216-220.

[Sterman 2000]
Sterman, J. *Business Dynamics: System Thinking and Modeling for a Complex World.* McGraw-Hill, 2000.

[Stokes 1994]
Stokes, A. & Kite, K. *Flight Stress*. Ashgate, 1994.

[Venkatesh 2003]
Venkatesh, V.; Morris, M.; Davis, G. B.; & Davis, F. D. "User Acceptance of Information Technology: Toward a Unified View." *MIS Quarterly 27,* 3 (Sept. 2003): 425-478.

[Vishwanath 2011]
Vishwanath, A.; Herath, T.; Chen, R.; Wang, J.; & Rao, H. "Why Do People Get Phished? Testing Individual Differences in Phishing Vulnerability within an Integrated, Information Processing Model." *Decision Support Systems, 51,* 3 (June 2011): 576-586.

[Wachtel 1968]
Wachtel, P. L. "Anxiety, Attention and Coping with Threat." *Journal of Abnormal Psychology 73,* 2 (April 1968): 137-143.

[Watters 2009]
Watters, P. A. "Why Do Users Trust The Wrong Messages? A Behavioural Model of Phishing," 1-7. *2009 eCrime Researchers Summit, eCRIME '09*. Tacoma, WA, Oct. 2009. IEEE, 2009.

[Weiner 2008]
Weiner, I. & Greene, R. *Handbook of Personality Assessment.* John Wiley & Sons, 2008.

[Whitten 1999]
Whitten, A. & Tygar, J. D. "Why Johnny Can't Encrypt: A Usability Evaluation of PGP 5.0," 169-184. *Proceedings of the 8th USENIX Security Symposium*. Washington, DC, Aug. 1999. The USENIX Association, 1999.

[Workman 2008]
Workman, M. "Wisecrackers: A Theory-Grounded Investigation of Phishing and Pretext Social Engineering Threats to Information Security." *Journal of the American Society of Information Science and Technology 59,* 4 (Feb. 2008): 662-674.

[Zald 2008]
Zald, D. H.; Cowan, R. L.; Riccardi, P.; Baldwin, R. M.; Ansari, M. S.; Li, R.; Shelby, E. S.; Smith, C. E.; McHugo, M.; & Kessler, R. M. "Midbrain Dopamine Receptor Availability Is Inversely Associated with Novelty-Seeking Traits in Humans." *The Journal of Neuroscience 28*, 53 (Dec. 2008): 14372-14378.

[Zurko 1996]
Zurko, M. E. & Simon, R. T. "User-Centered Security," 27-33. *Proceedings of the ACM New Security Paradigms Workshop*. Lake Arrowhead, CA, Sep. 1996. ACM, 1996.

[Zurko 2005]
Zurko, M. E. "User-Centered Security: Stepping Up to the Grand Challenge," 187-202. *Proceedings of the 21st Annual Computer Security Applications Conference (ACSAC'05)*. Tucson, AZ, Dec. 2005. IEEE, 2005.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
(Leave Blank)	January 2014	Final

4. TITLE AND SUBTITLE	5. FUNDING NUMBERS
Unintentional Insider Threats: Social Engineering	FA8721-05-C-0003

6. AUTHOR(S)

The CERT® Insider Threat Center

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)	8. PERFORMING ORGANIZATION REPORT NUMBER
Software Engineering Institute Carnegie Mellon University Pittsburgh, PA 15213	CMU/SEI-2013-TN-024

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSORING/MONITORING AGENCY REPORT NUMBER
AFLCMC/PZE/Hanscom Enterprise Acquisition Division 20 Schilling Circle Building 1305 Hanscom AFB, MA 01731-2116	n/a

11. SUPPLEMENTARY NOTES

12A DISTRIBUTION/AVAILABILITY STATEMENT	12B DISTRIBUTION CODE
Unclassified/Unlimited, DTIC, NTIS	

13. ABSTRACT (MAXIMUM 200 WORDS)

The research documented in this report seeks to advance the understanding of the unintentional insider threat (UIT) that derives from social engineering. The goals of this research are to collect data on additional UIT social engineering incidents to build a set of cases for the Management and Education of the Risk of Insider Threat (MERIT) database and to analyze such cases to identify possible behavioral and technical patterns and precursors. The authors hope that this research will inform future research and development of UIT mitigation strategies.

14. SUBJECT TERMS	15. NUMBER OF PAGES
unintentional, insider, threat, human, factors, decision, risk, mitigation, social, engineering, phishing, cybersecurity	109

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UL