Spy the Lie
Detecting Malicious Insiders
March 2018
CF Noonan
DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor Battelle Memorial Institute, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or Battelle Memorial Institute. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

PACIFIC NORTHWEST NATIONAL LABORATORY
operated by
BATTTELLE
for the
UNITED STATES DEPARTMENT OF ENERGY
under Contract DE-AC05-76RL01830

Printed in the United States of America

Available to DOE and DOE contractors from the
Office of Scientific and Technical Information,
P.O. Box 62, Oak Ridge, TN 37831-0062:
ph: (865) 576-8001
fax: (865) 576-5728
e-mail: reports@orfield.osti.gov

Available to the public from the National Technical Information Service
501 Shawnee Rd., Alexandria, VA 22312
ph: (800) 553-NTIS (6847)
e-mail: orders@ntis.gov <http://www.ntis.gov/about/form.aspx>
Online ordering: http://www.ntis.gov

This document was printed on recycled paper.
8/2010
Spy the Lie

CF Noonan

March 2018

Prepared for
the U.S. Department of Energy
under Contract DE-AC05-76RL01830

Pacific Northwest National Laboratory
Richland, Washington 99352
Abstract

Insider threat is a hard problem. There is no ground truth, there are innumerable variables, and the data is sparse. The types of crimes and abuses associated with insider threats are significant; the most serious include espionage, sabotage, terrorism, embezzlement, extortion, bribery, and corruption. Malicious activities include an even broader range of exploits, such as negligent use of classified data, fraud, cybercrime, unauthorized access to sensitive information, and illicit communications with unauthorized recipients. Inadvertent action or inaction without malicious intent (e.g., disposing of sensitive documents incorrectly) can also cause harm to an organization. This literature review paper will explore insider threat, specifically behaviors, beliefs, and current debates within the field. Additionally particular focus is given to deception, a significant behavioral component of the malicious insider. Finally, research and policy implications for law enforcement and the intelligence community are addressed. The cut-off date for literature reviewed for this report is July 2016.
Figures

Figure 1. Typology of deviant workplace behavior .................................................................3.30
Figure 2. The critical pathway to insider threat behavior .......................................................4.35
Tables

Table 1. Psychosocial Indicators of Insider Threat .................................................................2.12
Table 2. Motives for deception ...............................................................................................3.15
Table 3. Linguistic cues useful for testing deception models ................................................3.21
Table 4. Verbal techniques used by deceivers during an interview or interrogation ..........3.23
1.0 Introduction

All organizations face security risks. In 2011, President Obama issued Executive Order 13587 – Structural Reforms to Improve the Security of Classified Networks and the Responsible Sharing and Safeguarding of Classified Information.¹ This was quickly followed in 2012 by the National Insider Threat Policy and Minimum Standards for Executive Branch Insider Threat Programs.² The Order and the Policy provide additional guidance for the development of insider threat programs in federal agencies that handle classified information. Despite the formulation of an insider threat strategy and policy, federal agencies nonetheless currently still grapple with how to implement their own programs and in the process, better understand the human and organizational issues surrounding the insider threat problem.

Insider threat is a hard problem; there is no ground truth, innumerable variables, and sparse data. We often fail to acknowledge the impacts of socio-cultural and organizational influences on a person’s capability, motivation, and opportunity to commit an insider crime. Edward Snowden is the most discussed public face of insider threat, who in 2013 leaked thousands of classified documents. Among other monikers, Snowden has been called a traitor, whistleblower, and patriot. Studies reveal younger generations (e.g. the Millennials¹) believe him to be a political hero. Despite being charged with espionage, global polls conducted by the ACLU reveal that Millennials have a positive opinion of Snowden and view his actions favorably.² This cult-hero status poses a significant challenge to agencies managing classified data. Studies show employees who feel valued and who identify with their employer’s mission/vision work harder and maintain a firm workplace attachment.³,⁴ The attachments made between an employee and

¹ Millennials are generally agreed to have been born between the mid-1980s and the late 1990s, however, some extend this date range through the year 2000.
² For more information see: https://www.aclu.org/blog/speak-freely/generation-snowden
their employer is referred to as organizational commitment.\textsuperscript{4, 5} However, the opposite can also happen. When an employee does not feel valued, they can exhibit a range of counterproductive work behaviors – everything from acts of violence to something as benign as calling in sick when not ill. These negative or maladaptive behaviors are used by employees against an organization and its representatives to exercise punishment (i.e., revenge) for perceived injustice.\textsuperscript{6, 7} A sizeable body of research has established the tie between individual differences in personality and workplace incivility when the individual is under stress. Specific volitional acts include aggression, deceit, hostility, sabotage, and theft among others.\textsuperscript{8, 9} This work will explore insider threat, specifically behaviors, beliefs, and current debates within the field. Additionally particular focus is given to deception, a significant behavioral component of the malicious insider.
2.0 Insider Threat, Espionage and Spying

The insider threat refers to harmful acts that trusted insiders might carry out—for example, something that causes harm to the organization or an unauthorized act that benefits the insider. The insider is generally referred to as “a current or former employee, contractor, or business partner who has or had authorized access to an organization’s network, system, or data and has intentionally exceeded or intentionally used that access in a manner that negatively affected the confidentiality, integrity, or availability of the organization’s information or information systems.” Insiders are trusted individuals. They have been vetted by an organization and in many cases have been granted special privileges such as access to data, facilities, equipment and materials. Trust is established through a variety of mechanisms including a background check, security clearance investigation, credit checks, personal interviews, drug tests, polygraphs, and more.

The insider threat is manifested when human behaviors depart from established policies, regardless of whether it results from malice or disregard. The types of crimes and abuses associated with insider threats are significant; the most serious include espionage, sabotage, terrorism, embezzlement, extortion, bribery, and corruption. Insider threat may also include intellectual property theft, purposely or unwittingly mishandling classified data, fraud, unauthorized access to sensitive information, and other forms of cyber offenses such as social engineering.

Insider threat can be defined as a “wicked problem;” there is no readily available quantitative data for the development of ground truth. Wicked problems are defined as social or policy issues that are difficult or impossible to solve which is due to numerous variables, sparse data, incomplete or contradictory information, and the interconnectedness of one problem with
another. In the case of insider threat, there is no one driver for why an individual resolves to commit acts of crime. The decision to engage in deviant behavior can be influenced by personal health predispositions, such as clinical depression, and manifestation of certain psychosocial behaviors. This is a multifaceted phenomenon encompassing personality, personal history, life events, and more. For all of these reasons, insider threat is difficult to model and therefore to predict.

By definition, cyber-crime is any crime committed via a computer network. Many insider abuses (e.g., unauthorized access to data, intellectual property theft) can be conducted via a computer as demonstrated in the recent classified data leakage cases involving Chelsea (née, Bradley) Manning, Edward Snowden, and most recently, Harold Martin III. Cyber security firms indicate insider threats are on the rise, are hard to detect, expensive to mitigate, and take the longest amount of time to resolve. Computer technology cannot fundamentally alter human behavior; rather, computers are a means to manipulate the physical world and facilitate malicious activity. Deployment of security technologies is not the only way to address the problem. Criminal and psychological profiling are approaches by which organizations can detect and decrease abuse from the inside. However, there are serious limitations in this approach as there is no one accepted profile of an insider.

2.1 Insider Threat Typology
The impact of insider threat activities largely depends on the type of activities an individual engages in and the intended outcome of those actions. Insiders vary greatly in their level of threat due to their role within an organization, their level of access (this includes computer access as well as physical access), and knowledge of resources (e.g., financial, intellectual property) and organizational policies/procedures. For example, a systems administrator has access to a larger
number of computing resources than does a data entry clerk. In addition to an employee’s role, access, and knowledge, insider activities can be influenced by people outside an organization. This includes disgruntled employees collaborating with a competitor or cybercriminal.

Researchers have extensively studied insiders’ motivation to attack.\textsuperscript{11, 18-24} As discovered in the literature there are two major types of insider threat – malicious/intentional and unintentional. Malicious intent refers to a desire to cause harm to an organization or its assets. Intentional insider threat abuses are significant; the most serious include espionage, sabotage, terrorism, embezzlement, extortion, bribery, and corruption.\textsuperscript{25} Malicious activities include an even broader range of exploits, such as negligent use of classified data, fraud, unauthorized access to sensitive information, data or materials, and illicit communications with unauthorized recipients (ibid.).

The unintentional insider inadvertently causes harm to the confidentiality, integrity or availability of organizational resources. While not the specific focus of this research, it is beneficial to understand the differences between willful disregard for authority (intentional and malicious insider threat) and simple negligence or ignorance of policies/procedures (unintentional insider threat). Inadvertent action or inaction through poor performance can also cause harm to an organization. Unintentional insider threat has been referred to in the literature by a number of terms including accidental, inadvertent, innocuous, and unintended insider threat.\textsuperscript{19, 26, 27} Examples of unintentional insider threat behavior are disposing of sensitive documents incorrectly or being the victim of social engineering.\textsuperscript{26}

2.2 Modeling, Motives and Insider Crimes

Insider theft is a crime and according to some criminologists, crimes are generally believed to be preventable and predictable.\textsuperscript{28} However, despite ample research into the psychology and
motivation of insiders who engage in actions that harm their employer, the fact remains that it is very difficult to predict insider exploits. Considerable anecdotal evidence and post-incarceration interviews suggest the possibility of detecting warning signs of impending insider attacks. Warning signs can include changes in behavior, verbal outbursts, negativity, security breaches, confrontational, and deceitful behavior.

Modeling and mitigating the insider threat has been explored for years and will not be discussed at length here. Major research groups such as the Software Engineering Institute at Carnegie Mellon University (CERT), various Universities, National Laboratories, law enforcement and intelligence agencies have all invested millions of dollars into understanding how to build a secure workforce and protect critical assets. Major forms of insider crimes are defined below.

The terms spying and espionage are used interchangeably and refer to the act of taking or obtaining information covertly and without permission of the owner. Industrial espionage is a form of espionage used to gather information about a company or organization and to acquire intellectual property or proprietary data/information. Competitive advantage can be achieved through the illegal acquisition of information. Espionage occurs in one of two ways – through the use of insiders or outright theft. Motivations for insiders vary and include experiencing a personal crisis (e.g., health, financial) or as an act of civil disobedience.

Employee behavior in the workplace intended to damage, disrupt, or subvert an organization’s operations for personal purposes are considered sabotage. Acts of sabotage can include delays in production lines, damage to property, and physical or financial harm to an employer, its employees, or its customers. Sabotage has occurred in almost every critical infrastructure sector and is generally believed to be prompted by several motives including
frustration and a feeling of powerlessness prompted by perceived organizational injustice. The goal of sabotage in some sense is to restore equity, to compensate the individual for a supposed grievance.

Deliberate deception to secure unlawful or unfair gain is referred to as fraud. Fraud can be committed through a variety of means including media, phone, in person, and via the internet. False representation, white collar crime, counterfeit prescription drugs, non-delivery of merchandise, identity theft and market manipulation are all forms of fraud. The motivation to commit fraud is typically financial or other personal gain (such as prestige). As pertains to insider threat, the most frequent target is personally identifiable information that can be obtained within the workplace and use of corporate credit cards for personal use.

Taking another person’s property without their permission is theft. Employee theft is widespread and occurs in all industries. Employers and employees vary in their definitions of theft and theft may be overlooked for various reasons including lack of policies dealing with the subject or fear of public perception. General theories on employee theft focus in two main theoretical camps – person theories and workplace theories. Person-based theories attempt to explain why some people would steal from an organization. The main areas of research focus on perceived need, deviant background, greed or temptation (opportunity), moral laxity (typically among younger staff), and marginality (e.g., low status/low rank). Work-based theories concentrate research on why specific organizations may suffer higher levels of theft and focus attention on three major characteristics; organizational climate, deterrence strategies (i.e., policies or lack thereof), and organizational fairness.

Individuals make a choice to engage in retaliatory behavior which can be rational or irrational. Identification of triggers and warning signs are often missed, overlooked or simply
ignored. Triggers are negative events, often unexpected, which influence decision-making and can create or intensify motive. Examples of triggers include abusive supervision, negative performance evaluations, being passed over for a promotion, etc. Warning signs include outbursts in the workplace, aggressive behavior, absenteeism, etc. In each of the above discussed expressions of insider threat (spying and espionage, sabotage, fraud and theft) in the workplace, characteristics of an individual, organizational culture, and the social context in which they live and work play a key role in one’s motive, intent, and opportunity to commit crime.

2.3 Current Research and Debates on Insider Threat

Practitioners in fields as diverse as criminology and computer science debate the topic of insider threat. The variety of literature on the topic ranges from predictive modeling to employee monitoring programs, psychological evaluations, more robust background checks, and technological monitoring solutions. All agree that the impacts of insider threat can be financially and politically devastating. Frequency of insider attacks is on the rise, with 62% of security professionals reporting increase in the number of insider threats in 2014-2015 and 40% expecting a data breach by malicious insiders this year.14 In addition, the social consequences of personality and the resultant socially influenced relationships have direct impact on work-related behaviors.39

It is a widely-held belief that anyone has the potential to become an insider threat, given the right combination of internal and external pressures, opportunity, and ability to commit the crime. Understanding what goes on in the mind of an insider has traditionally been relegated a secondary role to that of investigating technical controls to prevent data leakage, etc.27 Research on psychology and motivation of insiders has been sporadic since the end of the Cold War. Motivation, opportunity and means are all required in order for an insider crime to occur. A well-
The known acronym for explaining underlying motives of insiders is MICE: money, ideology, coercion and ego. The four MICE drivers are often said to lie at the heart of why insiders betray their employers or state secrets. Insiders don’t commit crimes because they can (ability); they commit them because they want to. Hence, motive precedes opportunity.

Research on personality’s influence on insider threat is gaining popularity. Psychometric tools such as the Myers-Briggs Type Indicator and Social Styles are commonly used in the workplace. However, they are criticized for poor validity and poor reliability. Personality inventories such as the Big Five and the Dark Triad are heavily used in insider crime studies. Recent research in psychology, organizational behavior, biosocial criminology, and linguistics all contribute to the growing body of knowledge on the study of cybercrime and insider threat. However, many of these approaches have far-reaching social, ethical and legal ramifications.

### 2.4 The Big Five

According to the Five Factor Model (FFM), there are five dimensions of personality traits that account for all individual differences which can be attributed to genetic and non-shared environmental factors. Over fifty years of personality research, including multiple studies, confirms the robustness of this model. The five factors are – neuroticism, extraversion, and openness to experience, agreeableness, and conscientiousness. Neuroticism is expressed by a negative nature. Individuals who score high in neuroticism are believed to self-select themselves into situations which foster negative emotion. Extraverts, on the other hand, are predisposed to experience positive emotions. Individuals with extraverted tendencies tend to have more friends and spend more time in social situations. Openness to experience is related to creativity, divergent thinking, low religiosity and political liberalism. Agreeableness is characterized by altruism, nurturance, higher levels of intimacy, caring and emotional support. And finally,
conscientiousness implies stronger work associations and rewards. Four of the five traits (altruism, extraversion, agreeableness, and conscientiousness) have positive relationships with overall job satisfaction. Neuroticism has been identified as the primary source of negative affectivity in the workplace. Openness to Experience predisposes individuals to feel both the good and the bad more deeply, making its relationship to overall job satisfaction unclear.

Each of the five factors contains six constituent facets (i.e., narrow traits), meaning that six facets make up each of the five factors. As example, Extraversion contains facets related to friendliness as well as facets related to dominance and energy such as thrill-seeking and assertiveness which are also associated with deviance. While the factors present on a spectrum – positive to negative – research tends to neglect the narrow traits in the prediction of workplace deviance. While the FFM has been around for a long time it is not without reproach and has received criticism for failing to address all individual differences in personality, particularly antisocial behavior.

2.5 The Dark Triad

Dark personalities are those individuals characterized by socially offensive traits which fall into the “normal” range. The Dark Triad originally consisted of three traits – narcissism, Machiavellianism, and psychopathy. Recently a fourth trait, sadism, has been added leading to a new moniker, the Dark Tetrad. Narcissists crave attention and are grand self-promoters. Machiavellians are viewed as callous calculated manipulators. Classic psychopaths are the most malevolent, with low empathy and anxiety. In addition, they are reckless, thrill-seek impulsively, and lie for immediate rewards. Sadists may physically or verbally abuse others; the bully or Internet troll exhibiting predatory behaviors are classic examples. It is recognized that there are overlaps in key features of the four personalities, most importantly that of callousness, or lack of
empathy towards others. Similar to the FFM, several questionnaires, scales, or surveys are used to test and evaluate individual responses. Dark personalities are almost always linked to deviant behavior which will be discussed in more depth later.

2.6 Psychosocial Indicators

Various demographic, behavioral and psychological indicators reveal data about an individual which may provide insight into whether or not one has the tendency to commit crime. Studies in workplace aggression, entitlement, counterproductive work behavior, Internet use, computer-use deviance in the workplace, information security, and criminal profiling all reveal psychosocial indicators of vulnerability or risk to insider crime behavior and/or tendencies. Behavioral precursors to IT sabotage cases include missing work, arriving late; leaving early, decline in job performance, etc. The Defense Personnel and Security Research Center conducts research on Department of Defense personnel suitability, security and policy. This includes, but is not limited to, espionage, cyber-psychology, insider threat, and personality disorders. Their findings indicate that personality disorders and vulnerabilities within organizations and industries impact the psychosocial behaviors of insiders. A recent study conducted with human resources professionals identified twelve leading psychosocial indicators (see Table 1) of behaviors that would cause enough alarm to elevate personnel monitoring. The study further revealed that low-risk indicators only became alarming when presented with one or more high-risk indicators. This study is important in that it illuminates the need for collection of multiple data types to inform risk mitigation processes. Additionally, this work reveals interesting parallels between descriptions of those who commit various insider crimes (e.g., theft, sabotage), the FFM and the Dark Tetrad. For example, self-centeredness is associated with narcissism from the Dark Tetrad.
Table 1. Psychosocial Indicators of Insider Threat. High-risk indicators are underlined for emphasis (Greitzer, et al., 2010)

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Disgruntlement</td>
<td>Employee observed to be dissatisfied in current position; chronic indications of discontent, such as strong negative feelings about being passed over for a promotion or being underpaid, undervalued; may have a poor fit with current job.</td>
</tr>
<tr>
<td>Not Accepting</td>
<td>Employee is observed to have a difficult time accepting criticism, tends to take criticism personally or becomes defensive when message is delivered. Employee has been observed being unwilling to acknowledge errors; or admitting to mistakes; may attempt to cover up errors through lying or deceit.</td>
</tr>
<tr>
<td>Feedback</td>
<td></td>
</tr>
<tr>
<td>Anger Management</td>
<td>The employee often allows anger to get pent up inside; employee has trouble managing lingering emotional feelings of anger or rage. Holds strong grudges.</td>
</tr>
<tr>
<td>Issues</td>
<td></td>
</tr>
<tr>
<td>Disengagement</td>
<td>The employee keeps to self, is detached, withdrawn and tends not to interact with individuals or groups; avoids meetings.</td>
</tr>
<tr>
<td>Disregard for</td>
<td>The employee disregards rules, authority or policies. Employee feels above the rules or that they only apply to others.</td>
</tr>
<tr>
<td>Authority</td>
<td></td>
</tr>
<tr>
<td>Performance</td>
<td>The employee has received a corrective action (below expectation performance review, verbal warning, written reprimand, suspension, termination) based on poor performance.</td>
</tr>
<tr>
<td>Stress</td>
<td>The employee appears to be under physical, mental, or emotional strain or tension that he/she has difficulty handling.</td>
</tr>
<tr>
<td>Confrontational</td>
<td>Employee exhibits argumentative or aggressive behavior or is involved in bullying or intimidation.</td>
</tr>
<tr>
<td>Behavior</td>
<td></td>
</tr>
<tr>
<td>Personal Issues</td>
<td>Employee has difficulty keeping personal issues separate from work, and these issues interfere with work.</td>
</tr>
<tr>
<td>Self-Centeredness</td>
<td>The employee disregards needs or wishes of others, concerned primarily with own interests and welfare.</td>
</tr>
<tr>
<td>Lack of Dependability</td>
<td>Employee is unable to keep commitments/promises; unworthy of trust.</td>
</tr>
<tr>
<td>Absenteeism</td>
<td>Employee has exhibited chronic unexplained absenteeism.</td>
</tr>
</tbody>
</table>

2.7 Ethical and Legal Considerations

While the FFM and Dark Triad psychometric tests provide personality dimensions measured with high reliability and validity, these tools are not without problems. First and foremost, the empirical study of antisocial behaviors is fraught with ethical complexities. For example, when conducting experiments on sadist behavior it is inappropriate to have individuals harm others so
alternative means of study must be developed. Secondly, there are extensive time commitments required to perform comprehensive personality inventories on individuals – and very few individuals have been studied longitudinally. This has led to the development of short personality inventories which are less costly to administer. That being said, quantitative evidence fails to support individuals who exhibit the categorical psychological concepts described in both the Dark Tetrad and the FFM as clinically psychopathological. In other words, attitudes and personality traits have poor predictive validity on future expressions of behavior. While narcissists can be extremely attractive during a job interview, downstream they are highly likely to exhibit counterproductive work behaviors such as white collar crime or intellectual property theft.

Recently, several lines of inquiry have been identified for future research on the social consequences of personality which are influenced by organizational/workplace factors, individual goals and values, and cultural components. Research shows that certain behaviors and disorders are influenced by genetic factors including antisocial personality disorder, impulsivity, and more. Biological or environmental factors (e.g., abusive or traumatic childhood experience) related to criminal behavior, specifically insider threat, is a nascent area of study which carries significant legal and ethical dilemmas. Stereotyping or profiling individuals who exhibit specific personality traits, neurological, or biological characteristics is not an acceptable means by which to discriminate in the workplace. They do, however, encourage employers to be more cognizant of acknowledging individual differences and addressing issues/situations before they escalate.

Current practice for insider threat is largely reactive – dealing with problems after they are known to have taken place. In order to move towards a proactive approach for mitigating the
insider threat, organizations will have to balance technological solutions with effective monitoring of employees.\textsuperscript{60} This negotiation of cyber/physical security and modeling of human behavior follows an approach advocated in the early 2000s\textsuperscript{61} for the identification of clues (technical and social) to predict, detect and then interdict insider attacks. Computer-based monitoring technologies\textsuperscript{62} such as tracking web surfing behavior, remote desktop viewing to track job performance, and monitoring personal email opened on work computers can be viewed as profiling which is a highly sensitive topic. Access to confidential conversations, medical records, and remote desktop viewing may not only be legally questionable but have the potential to adversely affect or undermine employee morale when discovered. It could be argued that acknowledgement of such tools in the workplace may not serve as a deterrent to insider crimes, they may either exacerbate problems due to lack of trust or could immobilize staff for fear of organizational retribution. Either way, staff productivity and attitudes about security are bound to decline. Attitudes are precursors to intentional behavior. Strong negative emotion regarding security practices can lead to counterproductive or even subversive workplace behaviors ranging from apathy to absenteeism and retaliation, feelings of vulnerability, violation and shame, and deep seated feelings of alienation.\textsuperscript{63-65} It is therefore paramount that employee attitudes are important to understand. No amount of employee monitoring will deter highly motivated, narcissistic insiders who believe they are invincible and are excellent at covering their tracks through manipulation of others. This highlights a complementary problem, detecting when employees are intentionally dishonest.
3.0 Deception Detection

Like insider threat, there are many definitions of deception. Deception has been defined as a way “to intentionally cause another person to have or continue to have a false belief that is truly believed to be false by the person intentionally causing the false belief by bringing about evidence on the basis of which the other person has or continues to have that false belief.” This definition focuses on active intent which specifically excludes unintended actions which might lead to deception. Deception is an intentional act that occurs when a communicator tries to control information in a message meant to evoke a particular response or effect. In light of these definitions, deception can take a variety of forms ranging from explicit fabrication to half-truths, vagueness, and concealment. Deception occurs for a variety of reasons. Motives for deception are generally characterized as self-, partner- or relationship-focused. Table 2 identifies the most common motives reported.

Table 2. Motives for deception (Burgoon, Guerrero, and Floyd, 2010, p. 407)

<table>
<thead>
<tr>
<th>I. Self-focused motives</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Protect, retain or gain resources</td>
</tr>
<tr>
<td>B. Ensure continuation of rewards or services from target</td>
</tr>
<tr>
<td>C. Protect or enhance self-image and self-esteem</td>
</tr>
<tr>
<td>D. Avoid abuse, conflict, punishment, or negative repercussions from target</td>
</tr>
<tr>
<td>E. Maintain privacy/avoid disclosure of secrets and risky information</td>
</tr>
<tr>
<td>F. Control conversational direction, length, or termination</td>
</tr>
<tr>
<td>II. Partner-focused motives</td>
</tr>
<tr>
<td>A. Protect target’s image or self-esteem</td>
</tr>
<tr>
<td>B. Protect target’s mental and emotional state (e.g., avoiding worry, hurt, fear, embarrassment)</td>
</tr>
<tr>
<td>C. Protect target’s physical state</td>
</tr>
<tr>
<td>D. Protect target’s relationship with third party</td>
</tr>
<tr>
<td>III. Relationship-focused motives</td>
</tr>
<tr>
<td>A. Avoid conflict</td>
</tr>
<tr>
<td>B. Avoid relational trauma</td>
</tr>
<tr>
<td>C. Avoid unpleasant, repetitive episodes</td>
</tr>
<tr>
<td>D. Avoid violation of role expectations</td>
</tr>
<tr>
<td>E. Avoid relational breakup</td>
</tr>
<tr>
<td>F. Obligatory acceptance</td>
</tr>
</tbody>
</table>


Deception is a part of everyday life. People tell an average of one to two lies per day, either in spoken or written form.\textsuperscript{70} Lying is a deliberate, conscious behavior, which many speculate leaves a trace or signal in its wake.\textsuperscript{71} And, despite the frequency with which we are exposed to lies, people’s ability to discriminate lies from truth is equal to that of chance.\textsuperscript{72}

The scientific study of deception has been around since the late 1800s. Deception is an intentional act, a deliberate attempt to mislead that occurs when a communicator tries to control information in a message meant to evoke a particular response or effect.\textsuperscript{73, 74} In light of these definitions, deception can take a variety of forms ranging from explicit fabrication to half-truths, vagueness, equivocation and concealment.\textsuperscript{68} Deception is interpersonal and includes practical jokes, forgery, imposture, consumer fraud, military and strategic deception, white lies, scams, hoaxes and more.\textsuperscript{75} Early studies among very young children concluded that they are horrible at deception. However, by the time children reach adolescence they achieve the requisite cognitive tools and capacity for intentional deception.\textsuperscript{75} There are two separate and distinct lines of inquiry in this area. The first includes research on polygraph testing and other physiological measurement tools. The second line of investigation focuses on behavioral cues, verbal and nonverbal, to detecting deception. Both areas of research are discussed.

The majority of studies in deception detection, regardless of approach or focus area, use human subjects in experimental settings. A recognized and significant limitation to research in this area is the focus on deception exhibited by a single suspect or deceiver.\textsuperscript{76} Humans are relatively poor performers at detecting lies and very little is known about social indicators of deception, though it is recognized that there are “unique cues to deception that may occur between co-conspirators or accomplices.”\textsuperscript{76} A pressing need in the law enforcement and intelligence communities is the development of tools and methods to collect reliable information.
during interviewing or screening of individuals and groups. There is a disadvantage as there has been very little research focused on deception by collusion among two or more individuals. This can include recruitment of human sources overseas, vetting of employees, or screening people at security checkpoints.

There are emotional, mental, and physical challenges to deception which can also play a role in its detection. Telling lies and engaging in deceptive behavior are physiologically more challenging than being truthful. Deceit is emotionally challenging as deceivers may experience fear and threat of being caught, guilt and shame of deceiving someone, or even elation at having been successful in deceiving someone. It is mentally challenging as deceivers need to create a believable story and commit it to memory, stick to it. And, deceit is physically challenging as deceivers try to control the physical signs of deceptive behavior (nervousness, eye movement, body language, etc.). Deception cues are grouped into three main categories: (1) visual (non-verbal) – any physical behavior, (2) vocal (para-verbal) – pitch of voice, tone/tension, and rhythm (number and length of pauses), (3) verbal - anything said or written. This research reveals effective deception detection isn’t just about what liars do but what they don’t do that matters. In other words, motivated liars manipulate their body language and speech patterns in ways that are designed to make them appear more truthful than they actually are.

### 3.1 Current Research and Debates on Deception Detection

The main concern with deception detection is that it is widely open for interpretation. The polygraph has a long and controversial history as a forensic tool and continues to be contested as an appropriate mechanism to judge an individual trustworthy of obtaining or maintaining a security clearance. And, while many states in the U.S. do not admit polygraph evidence, eighteen do but only if certain requirements are met. In past court cases, polygraph results were
considered too unreliable to be admitted as evidence and have been found to infringe upon jurors role as triers of facts. Additionally, research in the late 1990s identified only four studies using the most common polygraph test (i.e., the Common Question Test) were rigorous enough to be accepted and published in peer reviewed scientific journals. Organizations such as AntiPolygraph.org find polygraphy so misleading and unethical that they publish advice on how to defeat the procedure and widely educate the public on countermeasures to include behavioral and chart-recording manipulation.

Verbal cues to deception online are of significant interest for a variety of law enforcement and intelligence needs (e.g., recruitment and radicalization). Primary differences between in-person deceit and online deceit are difficult to tease apart. Deception occurs in a variety of settings and for a variety of purposes. Accordingly, the linguistic features relevant to one context do not necessarily hold in another context (e.g., in-person communication versus instant messaging). The majority of deception investigations has been conducted by researchers in the psychological domain and often relies upon laboratory or other controlled studies. However, more recent studies are addressing deception in online social media platforms and through social engineering.

There are very few studies that explicitly address deception in instant messaging (IM), discussion boards and other forms of computer mediated communications. When trying to bridge the gap from in-person, verbal communications to email and IM, we must understand how the online “modalities restrict, highlight or amplify certain [deception] cues and the ways that they impact the choices deceivers have available to them in managing their messages.” As example, deceivers using IM can monitor the interaction as it occurs, and are not burdened by linguistic cues (i.e., hedging) or nonverbal mechanics (i.e., fidgeting, averting eyes) that might otherwise
be incriminating. Research conducted at MITRE on what is coined “cyber-deception” includes everything from deceptive online advertising, falsified user profiles on dating websites, cyber-espionage, lying in email or via VoIP conversation, to manipulating online images, etc. 

Basically, any form of deception of information transmitted via the Internet.

Topics not investigated herein but which may be relevant to understanding deception detection and its application to the law enforcement and intelligence communities include:

- Pathological liars (how they may be detected and if there are any core linguistic traits);
- Cultural, gender and age based differences in the detection of deceit;
- Studies involving varying degrees of deceit – ‘white,’ ‘serious,’ and ‘high-stakes’ lies;
- And, the possible convergence of deception detection research with studies into misinformation or disinformation.

As mentioned previously, deceit at a group level is very rarely studied. Additional research on group deception may be valuable as it pertains to insider recruitment in discussion forums, the proliferation of deception, etc. This may also be relevant for identifying deflection or diffusion of responsibility for actions or events. Finally, it is widely known that the majority of deception detection research has been conducted on native English speakers (in particular, American English). It would therefore be valuable to identify (or conduct) specific studies on mixed-culture and mixed-native language groups to provide insight into this rather unbalanced area of inquiry.

### 3.2 Linguistic Cues to Deception

Because deception occurs so frequently and can be quite consequential, there is a considerable body of work devoted to identifying which linguistic cues are related to deception and how to classify deceptive and non-deceptive speech using machine learning techniques. While it is highly unlikely that scientific research will determine the one clue that without a doubt signifies deceit in verbal or written forms of communication, there are linguistic behaviors which provide
a window into detecting deception. For example, several published studies investigate measures of quantity which have included assessment of the number of words/clauses/sentences produced; speech rate, length of individual words, and number of words from particular grammatical categories such as noun or verb.\textsuperscript{74, 85} Other behaviors researched include expressivity, affect, causation, diversity, redundancy, informality and specificity.\textsuperscript{74, 85} In 1981, Zuckerman, et al. published the first comprehensive meta-analysis of cues to deception.\textsuperscript{73} Their search for all reports of the degree to which verbal and nonverbal cues occurred differentially during deceptive communications compared with truthful ones produced 159 estimates of 19 behavioral cues to deception from 36 independent samples. Two decades later DePaulo, et al. expanded this review to the results of 1,300 estimates of 158 cues to deception from 120 independent samples.\textsuperscript{73} While the majority of these focus on verbal behavior and non-verbal cues, many of the studies involving written deception are useful for linguistic analysis of deception in computer mediated communication. This work found that liars generally provide fewer details; make more negative statements; sound more uncertain, impersonal, evasive, and unclear; and produce more words that distance themselves from their statements and the person or people to whom they are lying when compared with truth-tellers.\textsuperscript{73} Hauch, et al\textsuperscript{85} concluded that liars use more words expressing negative emotions, positive emotions, more emotional words in general, more motion verbs, and more negation words. In contrast, truth-tellers make use of more self-references, other references, exclusive words, more tentative words, and time-related words than liars. There were no discernible differences between liars and truth-tellers in word count. In a recent publication researchers\textsuperscript{86} examined 30 verbal and non-verbal communication indicators previously used with automated linguistic analysis tools. Their work specifically analyzed linguistic cues extracted from 367 written statements prepared by suspects and victims of crimes on military bases and
determined that only seven linguistic cues (see Table 3) proved useful for testing models of deception constructs with real-world, high-stakes data (versus controlled laboratory experiments).

| Table 3. Linguistic cues useful for testing deception models (Fuller, et al., 2013) |
|---------------------------------|------------------------------------------------|
| **Linguistic Cue** | **Description** |
| Quantity | Number of words |
| Specificity | Words used which establish the context of the statement in space/time |
| Affect | Emotion |
| Diversity | Variation in language |
| Uncertainty | Avoidance of relevant information, providing generalizable data |
| Non-immediacy | Use of words to create psychological closeness or distance |
| Activation | Vividness or intensity of word choice |

The most frequently cited tools for analyzing linguistic content for deception are the Linguistic Inquiry and Word Count (LIWC) and Whissell’s Dictionary of Affect in Language. In addition, natural language processing algorithms such as Coh-Metrix are utilized. Coh-Metrix has over 700 indices of computed language characteristics that have been validated across a variety of psychological domains and is unique in that it tracks linguistic features based on cognitive and social factors that are hypothesized to influence deception. Recent work utilizing the LIWC revealed individuals with word pattern usage linked to Big Five personality traits associated with an increased risk of insider threat behavior.

Linguistic cues have been investigated from a law enforcement perspective. When individuals provide verbal accounts of events they include important linguistic and structural features. These features can be examined for veracity and deception indicators. Oral accounts are coupled with written statements and various analytic methods are employed. These include criteria-based content analysis, behavioral assessment, scientific content analysis, and many more. The major distinction between oral and written accounts is the latter is a more deliberate and conscious process. Written discourse is longer, has more complex structure, and follows a
linear pattern, requiring sequence and clear patterning. Oral accounts, on the other hand are often more free-form, more spontaneous and often change depending on situational cues including body language and other proxemics. Specific attributes of written text along with their frequency, use, and context have been identified as providing insight into truth or deceit.\(^8\) These include: equivocation, negation, prologue attributes, unique sensory details, emotions, and quoted discourse. A positive relationship between equivocation and deception are identified in oral statements and extend to written statements. The simple presence of negation may not be solely indicative of deception. The frequency, density and location of negation in a text may prove significant for determining the veracity or deception of a statement. Truthful, experienced memories contain more sensory information than do constructed memories. In addition, unusual details are strong indicators of veracity in oral statements. As far as emotions are concerned, their research found that “memories of experienced events included more affective information, such as emotional reactions, than did memories of created events,” and “that in fabricated accounts references to emotions might be omitted altogether.”\(^8\) Direct quotations also provide details not expected or typically provided in a fictitious or deceptive recounting of an event. Based on their work analyzing written statements produced by individuals who were party to police investigations in the United States, Adams and Jarvis determined that deceptive statements differ from truthful statements in both structure and content.\(^8\) In addition, former CIA officers identify several additional verbal techniques (see Table 4) used to indicate deception during an interview or interrogation.\(^9\) It is widely acknowledged that many of these techniques are used by truthful individuals; it is the context and number of verbal cues, or clustering, that will reveal a liar. Additionally, when linguistic cues are coupled with shifts in body language and facial expressions, the likelihood of identifying a liar increases dramatically.
Table 4. Verbal techniques used by deceivers during an interview or interrogation (Houston, Floyd, and Carnicero, 2012)\textsuperscript{90}

<table>
<thead>
<tr>
<th>Verbal Technique</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Failure to Answer</strong></td>
<td>Not responding to a direct question.</td>
</tr>
<tr>
<td><strong>Denial Problems</strong></td>
<td>Absence of explicit denial of something. Can be nonspecific denial or isolated delivery of denial.</td>
</tr>
<tr>
<td><strong>Reluctance or Refusal to Answer</strong></td>
<td>An expression used to dodge or evade responding to a question. Example: <em>I’m not sure I’m the right person to talk to. I’m not sure I can answer that.</em></td>
</tr>
<tr>
<td><strong>Repeating the Question</strong></td>
<td>Repeating the question provides an individual with the opportunity to think through their response, i.e., buy time.</td>
</tr>
<tr>
<td><strong>Non-Answer Statements</strong></td>
<td>Provides an individual with the opportunity to avoid or think through their response. Examples: <em>That’s a good question... I’m glad you asked that... I knew you were going to ask me that... That’s a legitimate concern.</em></td>
</tr>
<tr>
<td><strong>Inconsistent Statements</strong></td>
<td>Inability to keep a story straight.</td>
</tr>
<tr>
<td><strong>Attack Mode</strong></td>
<td>When facts of a situation put strain on an individual the stress can force them to attack. Typically the verbal response is coupled with other behavioral cues to deception.</td>
</tr>
<tr>
<td><strong>Inappropriate Questions</strong></td>
<td>Answering a question with a question that is unrelated.</td>
</tr>
<tr>
<td><strong>Overly Specific Answers</strong></td>
<td>Answering a question with too much technical detail or a very narrow focus.</td>
</tr>
<tr>
<td><strong>Inappropriate Level of Politeness</strong></td>
<td>Sudden and uncharacteristic increase in politeness when responding to a question with the purpose of increasing likeability.</td>
</tr>
<tr>
<td><strong>Inappropriate Level of Concern</strong></td>
<td>Attempts to diminish the importance of an issue. May even joke about the issue. Examples: <em>Why is this such a big deal? Why is everybody worried about that?</em></td>
</tr>
<tr>
<td><strong>Process or Procedural Complaints</strong></td>
<td>Offensive response to questioning. This may be a delay tactic, similar to repeating the question or non-answer statements. May deflect. Examples: <em>Why are you asking me? How long is this going to take?</em></td>
</tr>
<tr>
<td><strong>Failure to Understand a Simple Question</strong></td>
<td>Changing typical word use in response to a question to minimize the scope or magnitude of impact.</td>
</tr>
<tr>
<td><strong>Referral Statements</strong></td>
<td>Individual refers to a previous response to an answer. Example: <em>As I said during our last meeting...Like I told the guy who asked that last time...As we stated in our corporate filings.</em></td>
</tr>
<tr>
<td><strong>Invoking Religion</strong></td>
<td>Dressing up a lie. Examples: <em>As God is my witness...I swear on a stack of Bibles.</em></td>
</tr>
<tr>
<td><strong>Selective Memory</strong></td>
<td>Claiming no recollection of an event. Examples: <em>Not that I recall...As far as I know...To the best of my knowledge.</em></td>
</tr>
<tr>
<td><strong>Exclusion Qualifiers</strong></td>
<td>Enable an individual to withhold certain information and respond truthfully. Examples: <em>Not really...Fundamentally...Basically...Probably.</em></td>
</tr>
<tr>
<td><strong>Perception Qualifiers</strong></td>
<td>Enable an individual to enhance credibility. Examples: <em>Frankly...To be perfectly honest...Honestly.</em></td>
</tr>
</tbody>
</table>
3.3 Physiological Cues

In addition to linguistic cues, a wide range of techniques, including the polygraph have been developed to assist in the detection of liars. Traditional polygraphy uses an array of physiological measures like heart rate, blood pressure, and electrodermal response to detect deception. It should be noted that physiological responses measured by the polygraph are postulated to be associated with deception. When the polygraph is applied to populations not trained in countermeasures it works fairly well and can discriminate lying from truth telling at rates well above chance. However, this is simply not good enough when it comes to protecting intelligence and law enforcement assets.

History of the polygraph dates to the early 1920s; the machine was invented by a forensic psychiatrist named John Larson. It was Leonarde Keeler, a protégé of Larson’s, who patented the first portable polygraph which is the prototype for all modern polygraph machines. Keeler went on to establish the first polygraph unit within a police department and his device was the first purchased by the Federal Bureau of Investigation. In the late 1930s there was a theoretical split between those who focused on standardized and objective approaches to the polygraph interview and use of the polygraph as an interrogation device with an emphasis on interviewer interpretation of an examinee’s behavior. During World War II the polygraph was used by the government for screening German prisoners of war and was eventually implemented within the Manhattan Project facilities at Oak Ridge and elsewhere. In both instances criminal activity, security breaches, and unauthorized disclosures of classified information were uncovered through use of the polygraph. Throughout the 1960s and 70s more widespread use of the polygraph became standard within U.S. federal agencies as well as private industry. In 1983 President Reagan issued National Security Decision Directive 84 authorizing all federal agencies
to polygraph employees to identify leaks of classified data. Directive 84 prompted an Office of Technology Assessment review of scientific evidence and validity of the polygraph’s utility in a screening capacity. Within three months the Directive was rescinded. Since that time use of the polygraph in federal agencies has increased dramatically. But, debates about the accuracy and integrity of polygraphy resurfaced in 1999. This followed significant security breaches at Los Alamos National Laboratory which prompted Department of Energy Secretary Bill Richardson to order polygraph tests for all of the Laboratory’s nuclear weapons scientists. When it was acknowledged that the polygraph was used to identify Wen Ho Lee as a spy at Los Alamos, public debates prompted the Department of Energy to retain the National Academy of Sciences to conduct a scientific review of the use of the polygraph. Their findings echo those previously discussed – accuracy rates for the polygraph in detecting lying were found at rates above chance but far less than perfect. Regardless of their findings, the polygraph continues to be used in all federal and many local law enforcement agencies for criminal investigations and employee screening.

Several new areas of research in neurophysiological measures including functional magnetic resonance imaging (i.e., fMRI) have been explored to support traditional polygraphs, identify liars, detect terrorists, and reveal secret intentions. Findings from many of the predictive studies discussed by Monteleone reveal 70-90% accuracy which is comparable to the results of the traditional polygraph: well above chance, but definitely below perfection. Advances in magnetic resonance imaging and electroencephalography (EEG) can now measure changes in brain activity due to emotions and behavior. These measurements can allow scientists to associate brain activity to cognitive processes adding new perspectives to state of the

3.25
art neuroscience research. The use of new technologies prompts ethics and privacy concerns, similar to those of employee monitoring discussed in the insider threat section.

An additional physiological mechanism adapted from the field of computer vision and computer mediated communication is eye tracking. Traditionally used as a mechanism to test content placement on websites and in usability studies, eye tracking studies reveal behavioral indicators (such as interest and emotion) due to cognitive load that can be used to optimize sales and product awareness and to positively increase knowledge retention and recall. In the field of visual cognition, studies have revealed conclusive evidence that eye-blinking behavior is related to cognitive processes such as deception.\textsuperscript{98-100} Through research on eye movements, including fixations (looking at the same place for a period of time), saccades (rapid eye movements) and pupil dilation response (change in pupil size), new non-intrusive commercial products such as EyeDetect\textsuperscript{®} from Converus\textsuperscript{®} have been launched. They assert an 85% accuracy to detect deception and are advocating for use of the technology in terrorist identification and combating fraud, theft and bribery. Neuro-ID’s Neuro-Screen\textsuperscript{™} solution, which claims to identify suspicious behaviors based on an individual’s typing, scrolling or mouse movements is another example of a commercial product focused on physiological response to stress and cognitive load as exhibited on a computer or smartphone. Both tools are being marketed to private industry and government agencies as complements to the polygraph due to lower costs of administration and decreased time to interpret and adjudicate test results.

3.4 Other Behavioral Cues

Studies show that the most of the commonly acknowledged behavioral “cue” for detecting a liar is avoidance of eye contact. However, this and other cues such as fidgeting, face touching, clearing the throat, and speech rate are conducted by truth tellers and liars at a fairly equivalent
rate. Body language alone is therefore an entirely unreliable mechanism for the detection of deception.

The majority of human communication is nonverbal; some studies indicate more than 50% of all communication is nonverbal.\textsuperscript{101} This estimate varies depending upon the type of communication and its purpose, cultural variables, and to whom the communication is being addressed. Nonverbal communication includes several subcategories – haptics, kinesics, vocalics, chronemics, and proxemics.\textsuperscript{102} Nonverbal and verbal communication deception techniques are used in coordination during interviews or interrogations to determine guilt or innocence.

Haptics, or communication through touch, is a form of interpersonal communication that conveys physical intimacy or aversion through lack of touch; hence, touch can be positive or negative. Touch is also believed to enhance or intensify emotional displays from other nonverbal communication modalities. Deceptive persons may experience a spike in cognitive load initiated by lying. This flood of hormones triggers the fight-or-flight response leading to an increase in self-soothing activities such as hand-to-face contact.\textsuperscript{103}

Kinesics, or body language, is communicated through a variety of mechanisms including facial expressions, hand movements, manipulation of a body part or item of clothing (e.g., nail biting or rubbing hands together, playing with a button).\textsuperscript{104} Similarly to language, body language does not necessarily reveal deception. It may indicate nervousness or uncomfortableness with a situation, person, or line of questioning.

Paralanguage, or verbal style, includes rhythm, speed, volume, and pitch, which can aid in interpreting the accompanying body language.\textsuperscript{69} Rhythm and speed can reveal emotions such
as apprehension or confidence. Volume provides insight into what is being said with a whisper or a shout. A rise in pitch can indicate fear; a drop in pitch may indicate sadness.

Chronemics, or the study of time, may be useful in the study of deception, specifically behavioral indicators in the workplace. The way a person perceives time and the role time-based concepts play in their lives are learned behaviors. There are four basic psychological time perspectives – past, time-line, present and future. For example, individuals who are present oriented are characterized as having low risk aversion, and being pleasure-seekers who live for the moment. In extreme cases this impulsivity can influence behavior regulation leading to aggression, destruction of property, problems with authority and compulsive lying.

Proxemics, or the study of human use of space, is another form of nonverbal communication. Hall breaks space into four areas to differentiate between public, social, personal, and intimate areas. The comfortable distance between people in each area varies culturally. Violating space norms can be a sign of aggression, emotional distancing, etc.

The study of micro expressions, popularized by the hit television show Lie to Me, focuses on the detection of very brief, involuntary facial expressions which last a fraction of a second. Micro expressions convey feelings such as anger, contempt or disgust and are revealed when an individual is deliberately or unconsciously concealing emotion. Commercial products such as Emotient™ and Affectiva® detect emotion and analyze sentiment, translating facial expressions into actionable information which could potentially be used in conjunction with the polygraph, EEG, fMRI and more.

The interplay between verbal and nonverbal behaviors is significant. During human interactions, cognitive processes are expressed through language and the body; this enables us to see partial truths or outright deceit. Close relationships allow for recognition of truth, half-truths
and lies. However, work relationships or casual acquaintances may lack the familiarity required to identify deceit or other aberrant behavior.

### 3.5 Deviance and Betrayal

Deviance is expressed in many forms including crime, drug addiction, alcohol abuse, pathological lying, etc. Acts of deviance are relative; they differ depending upon the culture and context. In general, deviant behavior is a departure from the norm for a society or group. Established rules or norms in many cases are codified into law, demarcating the good from bad, legal from the illegal and distinguishing behaviors, beliefs, and characteristics that are valued from those that are disvalued.\(^{107}\)

Deviance is understood in relation to group expectations or norms, social tolerance, and sanctions.\(^{107, 108}\) Norms are expectations of behavior which, when followed, are tolerated by a society. Sanctions are practices of social control to keep deviance within acceptable limits and define ways in which to deal with violations. Positive sanctions are rewards for acceptable behavior whereas negative sanctions yield punishments for violation of social norms (e.g., time off without pay).

Workplace deviance, also referred to as counterproductive work behavior, refers to behavior which is designed to protest or voice dissatisfaction in the workplace and results in actions that harm organizations or employees.\(^{109}\) Psychological contract breach, or an employee’s belief that their organization has failed to fulfill terms of agreement believed to exist between the two parties, can lead to maladaptive, retaliatory or deviant behaviors to include theft of office supplies, disobeying direct instructions, wasting time, aggression towards co-workers, cyber loafing, and other forms of cyber deviancy.\(^{6, 53, 54}\) Employees who do not behave in a productive manner are deviating from work norms. Organizational deviance is defined as
“unsanctioned nonproductive use of an employer’s time or property.” In order to better understand the severity of deviant workplace behaviors, Robinson and Bennett\(^\text{12}\) provide a typology (see Figure 1) which addresses a range of organizational and inter-personal deviance. While not exhaustive, the quadrant provides a robust framework for studying a range of workplace offenses.

**Figure 1.** Typology of deviant workplace behavior (adapted from Robinson and Bennett, 1995)\(^\text{12}\)

Acts of deviance and betrayal of workplace norms are often retaliatory and overt. The spectrum of these offenses may be minimal like practical jokes to the extreme such as sabotaging a production line. Studies suggest individuals with certain personality dispositions are more likely to engage in antisocial behaviors or to direct harmful actions against other people, groups, organizations or entities. The specific personality traits said to play a critical role in counterproductive work behaviors include narcissism, neuroticism, anxiety, anger, and the need
to be in control.9,110 This parallels recent research27,31,87 on the relationship of word use to personality traits indicative of known psychosocial indicators of insider threat such as the Big Five and the Dark Triad.111,112 Individuals’ behaviors are simultaneously influenced by multiple personality traits, situational factors, and the interactions between the two. The ability to understand the determinants of deviance, identify employee predisposition to behavioral and linguistic indicators of insider threat, and detect shifts in provocative or malicious activities serves as a warning sign to provide the opportunity for workplace intervention strategies.
4.0 Discussion

This work explored insider threat, specifically behaviors, beliefs, and current debates within the field. In particular, detailed focus was given to deception, a significant behavioral component of the malicious insider. Historically, neither of the topics is new – insider threat and deceit have been taking place since time immemorial. Relatively recently, it has been publicly acknowledged that the federal agencies which deal with issues of national security lack robust programs or have serious deficiencies. Mandates for U.S. government agencies to develop insider threat programs exist. However, determining how insider threat and deception countermeasures reduce these national security risks is difficult at best to ascertain.

Insider threat research has suffered from two main problems, limited data and validity of research conclusions. Modeling and simulation approaches from the social and behavioral sciences can be used to generate synthetic data and build in control groups and/or comparison groups. Additional modeling methodologies from military intelligence and counterterrorism have potential application to insider threat as well. Both of these approaches may contribute to the development of a scientific discipline of insider threat to rigorously test hypotheses, confirm or refute existing theory, or lead to the development of new theory.

Since computer networks are the mechanism by which insiders like Edward Snowden can steal secrets, there is a tendency for organizations to view insider threat as strictly a technology problem. This approach clearly negates the individual, organizational and industrial risks which influence insider threat. The variability in human motivation, intent and opportunity and its symbiotic relationship to the workplace makes solving the insider threat a ‘wicked problem.’

Insider threat is innately multidisciplinary. It requires the contributions of social scientists, statisticians, risk analysts, cyber security, physical and personnel security, human
resources, legal staff and policymakers to understand, model and mitigate insider crime. This divergence of scientific fields and varying organizational charters makes it difficult to “speak the same language.” As example, a recent article discussed more than 40 unique definitions for insider and insider threat. This lack of standardization is a noted problem for researchers. Does insider threat’s ‘wickedness’ encourage a broader definition of what has traditionally been considered insider threat? Some of the more recent definitions of insider threat have been expanded to include acts of workplace violence, suicide and homicide, specifically in industries like aviation and mass transit and the connection between insider threat and terrorism. Practitioners therefore, must be explicit when defining what an insider and insider threat are within their organization as this will have direct impact on insider threat mitigation strategies.

Empirical studies of insider acts, including espionage, demonstrate a comprehensive set of behavioral factors and patterns of individuals and the organizations in which they were affiliated. There are several risk factors identified in the literature including psychosocial/behavioral, organizational, and personal predisposition. Psychosocial and behavioral indicators include characteristics like difficulty accepting feedback, anger management issues, disengagement, disregard for authority, confrontational behavior, stress (work related, personal, financial, familial), among others. Organizational risk factors include complacency, things like allowing lapses in training, and a failure to follow and implement industry wide security policies and procedures. Familiarity can cause a manager to be blind to behavioral indicators exhibited by staff, choosing to ignore red flags rather than counsel the staff member. Internal politics distract management and administrators from the day to day work being conducted on their watch. In addition, the following factors can contribute to work-related stressors: downsizing and outsourcing, job insecurity, working long hours, and a poor work-life
Personal predispositions describe medical or psychiatric disorders, history of rule violations, maladaptive personality disorders (such as habitual lying) or a lack of social skills, and excessive travel, possibly indicative of divided loyalties. Placed into a critical path approach (see Figure 2), it is easy to see how a troubled employee’s behavior can escalate over time in response to personal predispositions and stressors/triggers. If risk signs are not identified and responded to by an employer, intervention strategies are likely to be inadequate or too late within the cycle to be effective. As pointed out by Shaw and Sellers, research in the field supports the assertions of the likelihood individuals will commit hostile acts against their organizations increases with the accumulation of various factors acting on them over a period of time. Mitigation or successful organizational interventions can remove individuals from the path towards attack. Thus, at any point on the critical pathway, an individual may exit from the path and not commit a crime. Many factors influence the likelihood of an individual staying on the path, including the organization’s policies and management practices that affect employee morale, economic/political influences, personal stressors (financial situation, medical crisis, family crisis, etc.). All individuals experience stressors, which can be both positive (e.g., marriage) and negative (e.g., divorce); the impact of stress is mitigated to the degree that an individual has access to resources (i.e., external resources such as money, social support, stable employment; and internal resources such as emotional stability, physical health, and emotional skills), but when stressors affect an individual who does not have adequate resources, distress occurs. Aberrant behavior (i.e., insider threat) is the individual’s response to chronic distress.
At the Federal level, the National Counterintelligence and Security Center at the Office of the Director of National Intelligence runs the National Insider Threat Task Force (NITTF). The NITTF is tasked with the development of a national insider threat program with supporting policy, standards, guidance and training. Earlier this year the NITTF published a Best Practices document for the purpose of providing information on insider threat awareness, case examples and nine key steps for program implementation. The Best Practices document identified key components for successful insider threat programs such as computer security, systems vulnerabilities, employee vetting, screening and training, and termination procedures. However, it lacks discussion of the critical path model and how to lessen the impact of co-worker deviance and incivility among employees. This is particularly important from a deterrence perspective. Future work in this area should include the development of employee education on the topics of social engineering, insider threat, and the identification of maladaptive
or counterproductive workplace behaviors and how to do so while maintaining workplace
privacy and ethics.

Within the workplace, coworkers may exert more influence over one another than do
their superiors. In fact, an employee is more strongly impacted by negative coworker behavior
than by positive coworker behavior. Deviant, dysfunctional or counterproductive work behaviors
conducted by one disgruntled employee may therefore have direct and lasting impacts on
colleagues. Workplace incivility can rapidly spiral out of control potentially leading to
increasingly aggressive behaviors.\textsuperscript{38, 48} Robinson, Wang and Kiewitz\textsuperscript{123} report that coworker
deviant behavior has been found to negatively impact others’ attitudes, affect and actions by (a)
direct impact (i.e., an employee is the target of a coworkers’ deviant behavior), by (b) vicarious
impact (i.e., an employee learns of or witnesses a coworkers’ deviant behavior), or by (c)
ambient impact (i.e., an employee works in an environment characterized by collective deviant
behavior).

Counterproductive workplace behaviors are crucial to identify, intercept, and alleviate.
Left unchecked workplace bullying, incivility, harassment, and verbal abuse can lead employees
to perceive workplace injustice and lessen their organizational commitment. Mistreatment of
coworkers has been shown to trigger emotional and behavioral responses directly related to
varying forms of workplace deviance including higher rates of absenteeism, theft, workplace
violence, and more.\textsuperscript{38, 123} In fact, Ambrose, Seabright, and Schminke\textsuperscript{11} found that injustice was
the most common cause of sabotage in the workplace. A brief review\textsuperscript{124} of life-course
development theories suggests that patterns of antisocial behavior, delinquency, crime,
psychopathy and impulsivity may be disproportionately represented in groups of individuals who
are referred to as having a “dual diagnosis” as being both victims and offenders of workplace
incivility. These findings clearly indicate the impact of organizational injustice on dysfunctional work behaviors.

In order to mitigate potential insider threat, organizations should match the severity of punishment to the perceived seriousness of the deviant act. For example, bullying and time card fraud may yield the same punishment; whereas terse emails and chronic tardiness should receive less severe punishment. Penalties should be standardized and applied across the enterprise. This ensures policy, rules, and regulations are perceived as fair and just by employees to the extent that similar behaviors are punished in a comparable fashion. Organizational commitment will be reinforced or may even increase if employees feel safe, cared for and treated respectfully by coworkers and supervisors alike.

4.1 Personnel Screening

As discussed in the body of the paper, background checks, security clearance investigations, credit checks, personal interviews, drug tests, and polygraphs are all used to some extent to screen current and prospective employees in law enforcement and intelligence positions who require access to classified data. Formalized adjudicative guidelines are used to evaluate each person careful weighing of a number of variables known as the whole-person concept. Available, reliable information about the person, past and present, favorable and unfavorable, are considered in reaching a security clearance determination.

Two new approaches are being explored as complements to the traditional personnel screening process; use of publically available social media information, and physiological deception detection tools. Security Executive Agent Directive 5 formally authorizes the use of social media by official investigators who are conducting background investigations for security clearances. Deception detection tools such as Converus® are under consideration as
enhancements to the polygraph. Application of both approaches within the personnel security regime will require higher levels of scrutiny to avoid undue infringements on privacy.

4.2 Privacy and Ethics

The implementations of insider threat and deception detection programs are not without serious privacy and ethical considerations. Privacy is generally characterized in terms of the rights of an individual or group to determine when, how, and to what extent information about them is collected and disseminated. Issues of security and privacy are founded in law and argued in the study of ethics. In the U.S., there is no single source of privacy law. For those working on government computing systems there is no expectation of privacy. And, in many organizations employees must consent to various forms of monitoring including remote viewing, email harvesting, review of access logs, and the like.

The debate regarding ethics is largely centered on the concept of trust. A key ingredient in employee morale is maintenance of the psychological contract (i.e., implied and unspoken mutual expectations) between employer and employee. When personal information is collected, even in the workplace, it places employees in a vulnerable position that relies on trust. This trust may be challenged with extensive organizational monitoring which may happen after a serious breach such as WikiLeaks. If employees do not feel due care and data safeguards are in place to protect their personally identifiable information, employee trust will be challenged.

4.3 Next-gen Workforce

Researchers generally agree that three generations dominate today’s workforce (i.e., Baby Boomers, Generation X, and Millennials). While labels and periods of year those labels encompass vary, Boomers were born between the early 1940s and the mid-1960s; Generation X were born between 1965 and 1979; Millennials were born between 1980 and 2000. As Baby
Boomers retire, Generation X (i.e., GenX) employees move up in the ranks and Millennials are being hired to backfill key positions. Law enforcement and intelligence agencies have serious challenges to negotiate a multi-generation workforce; these include work values, attitudes, motivation, motivation, organizational citizenship behaviors and more.128-132

Boomers are typically stereotyped as lifers or long-term employees, hardworking and optimistic. GenX is characterized by competitiveness, skepticism, informality, and desire for work-life balance. The youngest generation, the Millennials or Generation Y, is influenced by the Internet, 9/11, the effects of social media on events like the Arab Spring, collaboration and the post-2008 global economic crisis. This generation is also more globally oriented, has higher levels of narcissism, is more likely to take risks, seeks instant gratification, and is generally found to be more defensive when receiving criticism than previous generations.130 Ramifications of these characteristics include overall lower levels of organizational commitment, conformity, and formality in dress, speech, and social distancing. Most important for the intelligence community is a study conducted by the ACLU in 2015. The multi-country study revealed Millennials have a positive opinion of Snowden and believe his actions will lead to increased levels of personal privacy rights and protections.133

Research on generational stereotypes is not always consistent with observed workplace behaviors129 which has implications for training programs, development of multi-generational teams, work values and attitudes. Caution is issued for human resource managers to avoid treating employees as members of a specific generation. Individual differences are likely to play a more prominent role in workplace behaviors than generational differences. Greater flexibility in work hours, locations and work-life balance may better address needs and values of all employees regardless of generation.
5.0 Conclusion

This research explored the relationship between insider threat and detecting deceit in the workplace, two broad and interrelated topics. The impact of the malicious insider can be calculated by the severity of the event (e.g., theft of formula of new pharma product) and resultant economic ramifications (i.e., in real dollars and expected earnings) or impact on national security. Insider events may be categorized as malicious/intentional or unintentional. The type of insider event does not necessarily have direct bearing on the level of consequence. For example, the victim of an inadvertent social engineering attack may cause more financial damage to a company than an individual act of corporate espionage. Factors to consider when implementing insider threat programs are financial (i.e., cost to implement tools, policies), social (i.e., securing ethics and privacy), technological (i.e., firewalls, robust password requirements, etc.), and psychological (i.e., maintenance of the psychological contract between employer and employee).

Individuals commit acts of high-stakes lies, sabotage, treason, and cybercrime. While there are various internal and external factors which contribute to the detection of malicious insiders, at its core this is a human problem which must be addressed through robust research agendas in the human, cultural, behavioral, and social sciences. Eliminating the insider threat completely is impossible. Evidence-based approaches such as those reviewed in this paper can be coupled with effective personnel screening including robust background investigations, interviews, behavioral observations, technical monitoring and employee policies and training programs to lessen the likelihood of repeated Snowden-like affairs.
6.0 References


57. Shechter OG, Lang EL. Identifying Personality Disorders that are Security Risks: Field Test Results. PERSEREC; 2011.


