Psychological Vulnerabilities to Deception, for Use in Computer Security

Jim Yuill, Dorothy Denning, Fred Feer

Abstract: Vulnerability to deception is part of human nature, owing to fundamental limitations of the human mind. This vulnerability is exploited by con artists and scammers, but also by the military, intelligence, and law enforcement communities for the purposes of operational security, intelligence collection on adversaries, and undercover operations against organized crime. More recently, deception is being applied to computer security, for example, through the use of honeypots. This paper describes psychological vulnerabilities to deception and how they can be exploited to outwit computer hackers. The paper draws upon research in psychology and fraud, and the military and intelligence deception-literature.

1 Introduction

The military, intelligence, and law enforcement communities have long used deception for operational security, intelligence collection on adversaries, and undercover operations against organized crime. In recent years, deception has also offered a promising means for strengthening computer security through mechanisms such as honeypots. This paper describes psychological vulnerabilities to deception and how they can be used for computer security to defend against hackers. The paper draws upon research in psychology and fraud, and the military and intelligence literature on deception.

President Lincoln observed, “you can fool all of the people some of the time” [BW82]. Indeed, vulnerability to deception is a part of human nature, arising from fundamental limitations, or weaknesses, of the human mind [Heu81]. This paper addresses eleven such weaknesses, which fall into two broad categories: biases and impaired thinking.

Biases are human tendencies of erroneous perception or erroneous cognition (i.e., erroneous reasoning). An example of a perceptual bias is the human tendency to perceive that which is expected. An example of a cognitive bias is the human tendency to form generalizations with insufficient information. Exploiting a target’s biases can help ensure a deception is successful. Biases are statistically predictable in that one can expect humans to generally behave in a certain way. However, biases provide no guarantee that a particular person will behave in that way at any given time. Thus, when a deception operation depends on the target’s biases, the deception’s success cannot be entirely certain.

Impaired thinking refers to a variety of psychological influences that can weaken a person’s judgment or reasoning abilities. Moral vices such as greed, for instance, can lead to errors in judgment. In deception operations, one can attempt to induce impaired thinking, for example, by presenting a “limited time offer” that causes the deception target to act hastily and recklessly. However, as with biases, deceptions that exploit impaired thinking cannot be guaranteed to succeed.

Despite their limitations, deceptions that exploit biases and impaired thinking will be more likely to succeed than ones that do not. By understanding these psychological vulnerabilities to deception, the deception planner can take advantage of them, as opportunities arise.

The next three sections address perceptual biases, cognitive biases, and impaired thinking, respectively. In total, eleven psychological vulnerabilities to deception are presented. These are summarized in Table 1. A final section concludes.

This paper's treatment of biases is adapted primarily from Richards Heuer's research [Heu81]. Heuer was a senior CIA analyst, who applied psychology research on biases to military and intelligence deception. We adapted those parts of his work that seemed most useful for computer security. The paper's section on impaired thinking is drawn primarily from two books on fraud [San84, San94]. They are from a notorious publisher of books on felonious activity.

2 Perceptual biases

Human perception, and hence response to deception, is strongly influenced by expectations and desires. The following sub-sections explain the role of expectations in perception, present deception techniques that exploit these expectations, and show how the target's desires can be exploited for deception.

2.1 The role of expectations in perception

"The adversary is often the best source for opportunities to deceive . . . the preconceptions of the victim provide the most fertile ground for deception." USMC deception manual [USM89]

The mind can only process a small portion of the information it receives from the senses, e.g., sight and sound [Heu81]. To cope with the voluminous and complex information it receives, the mind constructs simplifying models of the world. Examples are social models that explain how people act and network models that characterize computer networks. These models are necessary for filtering the overwhelming information received from the senses. For example, when sniffing network traffic, the hacker's network model helps the hacker comprehend the voluminous data received.

One of the strongest influences on perception is one's *expectations*. There are several types, including *preconceptions*, *assumptions*, *mind sets*, and *stereotypes*. Expectations arise from diverse sources, such as past experience, training, and culture. Also, different circumstances evoke different sets of expectations. For instance, a hacker will reasonably expect different traffic on banking and university networks.

Expectations are necessary for perception. Correct expectations provide relevant and true perception. Wrong expectations can impair perception or cause irrelevant and false perception. Types of wrong expectations include *premature judgments* and *prejudices*.

In the military and intelligence literature, one of the primary deception principles is to exploit the deception target's expectations: in general, it is easiest to persuade the target to believe deceptions that are consistent with his expectations [Dew89, Heu81, JDD96, USM89]. A CIA deception study states it this way:

> *"It is generally easier to induce an opponent to maintain a preexisting belief than to present notional evidence to change that belief. Thus, it may be more fruitful to examine how an opponent's existing beliefs can be turned to advantage than to attempt to alter these views"* [CIA80].

In general, deceptions that are contrary to the target's expectations should be avoided, if possible [Heu81].

The target's expectations determine what things he notices and how he interprets them. In general, deceptions that are consistent with these expectations will be more readily received and believed. For instance, when hackers investigate a highly-secure network, they expect its vulnerabilities to be subtle and obscure, not glaring and obvious. These expectations can be exploited when building honeypots with vulnerable servers. The vulnerable servers will be more readily recognized and believed if they are consistent with the hackers' expectations.

In human perception, recognizing unexpected phenomenon requires more information, and more unambiguous information, than recognizing expected phenomenon [Heu81]. Thus, it is easier to build deceptions that are consistent with the target's expectations. Deceptions that deviate from these expectations must portray more information, and more unambiguous information, than deceptions that show what the target expects. For instance, when building a honeypot impersonation of a web server, it is better to put the honeypot on port 80 than on, say, port 22. This is because a hacker expects to see a web server on port 80, but not on port 22. If the hacker pings port 80 and gets a response, the hacker will assume it is a web server. Even though a honeypot could be placed on port 22, it will have to provide more information than a ping response to lead a hacker into believing that it is a web server.

Another aspect of expectations is that they are resistant to change [Heu81]. After a judgment about the essential characteristics of a thing are made, a person will continue to perceive it in the same manner even if the data are ambiguous. Further, once an expectation is formed, there is a tendency to assimilate new information in a manner consistent with the expectation. This tendency is greater the more ambiguous the new information and the more confidently the expectation is held [Heu81, Jer68]. Thus, when new information contradicts a person's expectations, the tendency will be to ignore or rationalize the information rather than to alter expectations.

Deception operations can benefit from the human tendency to resist changing one's expectations. Once the target has received and believed a deception, there is always a risk that the truth will leak out and reveal the deception. However, if the target is confident of his expectations, or if the leaked truth is ambiguous, then the target will likely reject such leaks and continue believing the deception [Heu81]. For instance, a hacker accesses a honeypot database-server on a company's intranet and believes it is a production system. When submitting queries to the database, the hacker notices extremely fast response times. Since he believes this is a production system, his expectations lead him to conclude that the server runs on a powerful computer. His expectations prevent him from realizing that the fast response times are due to him being the sole user of a honeypot.

2.2 Exploiting expectations

A target's expectations can be viewed along to two dimensions: whether they relate to his opponent or himself, and whether they relate to a course of action or to capabilities. The following describes the resulting four possibilities:

Exploiting the target's expectations regarding his opponent's course of action

One of the most effective techniques for exploiting expectations works as follows: if the target expects you to do A, then deceptively lead him to believe you are doing A, but do B instead [DH82b]. When doing the unexpected, the deception planner's task is to provide information that reinforces the target's expectations, while minimizing information that contradicts them. The power of expectations can cause the target to be an "unwitting but cooperative victim" in the deception.

To illustrate, a social-engineering technique used by hackers involves calling a system administrator and requesting an account and password. If the system administrator detects the con, he can deceptively exploit the hacker's expectations by providing an account and password for a honeypot that resembles the real system.

Exploiting the target's expectations regarding his opponent's capabilities

A common deceptive tactic is to portray weakness where one is strong, and strength where one is weak [USA88]. This deception can be simple to pull off when the target over-estimates his opponent's weaknesses. All the opponent need do is portray the weakness that the target expects. As an example, bullies always assume their victims are relatively weak, so a victim who is stronger can feign weakness, to his advantage.

In more general terms, a target's expectations include estimates of the opponent's capabilities. If the target underestimates or overestimates these capabilities, his false belief can be exploited. For example, a particular network has a highly effective intrusion detection system (IDS), and its capabilities exceed conventional IDSs. When hackers are detected and apprehended, the network's IDS capabilities can be kept secret by attributing detection to conventional IDSs, such as log files. Hackers will be vulnerable to this deception due to their expectation of conventional IDS capabilities.

Exploiting the target's expectations regarding his own course of action

The target's expectations can be exploited to deceptively manipulate his course of action. To induce the target to continue his current course of action, deception can portray favorable conditions that the target expects. To induce the target to change his course of action, deception can portray unfavorable conditions that the target considers possible or likely. For example, one of the primary uses of honeypots is collecting hacker intelligence. When hackers access the honeypot, hacking can be encouraged by deceptively portraying both what he expects and what he wants.

Exploiting the target's expectations regarding his own capabilities

The target can underestimate, or overestimate, his own capabilities. For example, a disgruntled employee believes he can safely attack his company's network from his home, and thereby avoid being identified. However, company officials, suspecting his malice, gave him a laptop with a hidden keystroke logger. The deceptive surveillance system will be aided by the target's expectation of security at home.

A limitation of exploiting target expectations is that, often, they cannot be known with adequate certainty. They reside in the target's mind, and they are subject to change. But expectations may be inferable [DH82b] from the target's capabilities and course of action. For example, a hacker's intelligence activity can reveal what he knows about a network, and, as a consequence, what he is likely to expect of it. In addition, the target's interactions with the external world set bounds on what he expects. For instance, hacking occurs within networks that use networking standards such as TCP/IP. These networking standards have predictable affects on hackers' expectations. In general, the target's personal predilections can be capricious and difficult to know, but his expectations of the external world can be known much more easily and reliably.

2.3 Exploiting desires

Besides expectations, a target's desires are an important, and exploitable, vulnerability. A CIA deception study quotes General Dudley Clarke, who led British deception operations in WWII, "all cover plans should be based on what the enemy himself not only believes but hopes for" [CIA80]. *Cover plans* are deceptions that hide true operations. Further, an authority on WWII British intelligence states that British deceptions "found their best targets in the obsessions of the enemy" [Wha69]. Also, in a paper on strategic military deception, Daniel and Herbig cite a study that found policy makers were vulnerable to "seeing what they devoutly wished to see, rather than what was there" [DH82b].

For deception operations to be successful, they must be received by the target and interpreted as intended. Then they must induce the desired action in the target. An effective way to accomplish this is to offer the target what he most desires. In Cliff Stoll's investigation of hackers who had penetrated a server at Lawrence Berkeley Labs, he discovered that the hackers were seeking information on nuclear weapons [Sto89]. So Stoll ran a sting operation, posting a deceptive file that stated where one could write to obtain such information. The hackers took the bait, and the sting operation's success revitalized the stalled investigation.

Although desires offer a valuable avenue for deception, they may play a less important role than expectations. According to Heuer, perception is influenced more by expectations than what one wants [Heu81].

3 Cognitive biases

We consider four types of cognitive biases. The first three are specific ways that people "jump to conclusions": the bias toward causal explanations, oversensitivity to consistency, and biases in estimating probabilities. The fourth bias relates to difficulties in detecting missing evidence. Psychology researchers have identified many other cognitive biases, but they are beyond the scope of this research. [1]

3.1 Bias toward causal explanations

There is a strong human tendency to seek causal explanations [Heu81]. However, causation is often not seen directly. Rather, it is perceived via a complex process of inference. In general, the process of forming causal explanations is subject to bias. The human desire to understand causation, for example, leads us to see order where it does not exist. Random things or events may wrongly be attributed to a non-existent cause, e.g. to purpose, design, or the effect of some orderly process. In addition, when observing the behavior of an organization, people tend to see the organization as more centralized, disciplined, and coordinated than it truly is [Jer68]. When people see only the outward actions of an organization, they tend to underestimate effects from internal problems and non-optimal processes.

[1] Dozens of biases are described in the Wikipedia entry "List of cognitive biases" (http://en.wikipedia.org/wiki/List_of_cognitive_biases).

Conspiracy theories typically exploit the bias toward causal explanations. In any large organization, there will be random mistakes, bad outcomes, and misbehavior among its members. The promoters of conspiracy theories can attribute these actions to the sinister schemes of the organization's leaders. Further, any missing evidence can be attributed to the conspirator's cleverness in hiding their schemes [Sch93].

In the domain of computer security, the deception planner can exploit the power of fallacious causal explanations and conspiracy theories by portraying fake security indicators. For example, a server can randomly generate ambiguous console messages that a suspicious hacker will attribute to detection of his activity. Legitimate users are instructed to simply ignore the messages. An example of such a message is:

[DEBUG #11] anomalous shell activity, generating IDS record at 13:43:02.36

The message is meant to be interpreted, by the hacker, as a debug statement that a developer accidentally left in an intrusion detection program. As another example, real systems can be given honeypot indicators, such as firewall rules that limit outgoing network traffic. From the hacker's perspective, the bogus indicators will be seen as confirming evidence of a honeypot. In both cases, missing indicators can be attributed to the network defender's stealthiness. The deceptive indicators also take advantage of hackers' hypersensitivity to detection, as described in Section 4.3 below. By exploiting these hacker vulnerabilities to deception, the false indicators can be random and ambiguous, and still be effective. This makes the deception easier to implement.

There are other ways that the bias toward causal explanations can be used to advantage. In situations where the deception target knows that deception is being used, this bias can cause him to see deception where it does not exist [Heu81]. When the target suspects deception, deception will be attractive as a causal explanation. If the evidence of deception is incomplete, the target can attribute the missing evidence to the deceiver's cleverness. For example, in World War II, there were a number of instances in which Ally plans fell into German hands [CIA80]. However, the Germans often disregarded the plans because they were thought to be deceptions. The Germans wrongly chose a causal explanation of deception, over the true explanation of Ally mistakes. Similarly, if hackers know a network uses deceptive security measures, then the hackers will likely view anomalous security mistakes as deceptive traps.

3.2 Oversensitivity to consistency

When evaluating information, people reasonably look for trends, patterns or other forms of consistency. However when there is consistency in small samples, there is a strong tendency to overestimate the relevance of the consistency [CIA80, Heu81]. The error lies in overlooking the inherent uncertainty of conclusions based on small samples. For example, in a study of psychology researchers, the researchers were observed to have "seriously incorrect notions about the amount of error and unreliability inherent in small samples of data" [TK71]. This bias is referred to as "the law of small numbers." For deception, a useful effect of the bias is that trends or patterns may be deceptively portrayed via a small amount of consistency, e.g., in operations or systems [Heu81].

Conditioning is a well-known deception technique, and it can take advantage of a target's oversensitivity to consistency. Conditioning works by deceptively portraying a particular pattern of operations, so that the target comes to expect that pattern [DH82b, JDD96, USA88]. Often, conditioning is used to create the comforting illusion that a standard operating procedure is being followed, so that the target will come to expect, and disregard, that operation. The ultimate purpose of conditioning is to exploit the false expectations that are induced in the target. The bias of oversensitivity to consistency can make it possible to condition targets quickly.

To illustrate the application of conditioning to computer security, consider a network that hides three valuable computers from hackers' scans by making the computers appear to be printers; that is, the computers' operating-system signatures look like printers. To further enhance the impersonation, the network's printers are all named after cities, e.g., Boston, so that hackers will be conditioned to associate computers named after cities with printers, after discovering a few printers. By naming the valuable computers also after cities, they are further hidden from conditioned hackers.

In addition to conditioning, there are other deceptions that can exploit a target's oversensitivity to consistency. One such deception is the exaggeration of computer security capabilities. For instance, over a short period of time, an organization publicly announces three incidents in which hackers were caught and prosecuted. The small sample would likely induce an exaggerated expectation of prosecution, among hackers.

3.3 Biases in estimating probabilities

Adversarial relationships are characterized by uncertainty. To cope with this uncertainty, opponents rely on probability estimates to aid decision making. These estimates, however, are vulnerable to the *availability bias*. It is the human tendency to overestimate things that can easily be imagined or recalled, and conversely, underestimate things than are not as easily imagined or recalled [Heu81]. How easily a thing can be imagined is influenced by many factors, such as how complex it is, and one's personal interests and degree of understanding. For example, it is relatively difficult to imagine things that are complex or foreign to our thinking, but they are not necessarily less likely. Also, how easily a thing can be remembered is influenced by factors such as how recently one has been exposed to it and how vivid the memories are. However, if something occurred recently, it does not necessarily mean it is more likely to occur in the future.

When a deception story is portrayed to a target who uses probability estimates to interpret the story, it may be possible to put the availability bias to work. For example, suppose a hacker makes probability estimates regarding the computers found during network scans. A honeynet can exploit the availability bias by portraying computers that the hacker can easily imagine or recall, such as a web server rather than a special-purpose machine.

During his famous hacking case, Cliff Stoll had to stop a hacker from downloading a particular file. However, he could not do this by unplugging the network cable, as that would alert the hacker to the surveillance [Sto89]. Instead, Stoll deceptively thwarted the download by jingling his keys across the communication line, thereby creating line noise that sporadically corrupted the data transfer. The hacker could only make speculative probability estimates about the communication problems. The deception would be aided by availability bias if it were consistent with normal network problems that the hacker had recently seen.

Since deception operations are hidden, the hacker who suspects deception must constantly assess the things he sees to determine if they are real. Such assessments typically involve probability estimates, e.g., "it is most probably a deception." The availability bias can help in exaggerating the use of deception when the target suspects deception. For example, to exaggerate a network's use of honeypots, deceptive honeypot indicators are placed on real computers. To help make the indicators believable, the network's real honeypots are widely publicized. The publicity places honeypots at the forefront of hackers minds, and thereby induces availability bias.

Humans are particularly vulnerable to availability bias when they conduct intelligence collection and analysis [Heu81]. They are looking for specific things and have rehearsed various scenarios in their minds. Having these things at the forefront of their mind is likely to bias their probability estimates when they encounter indicators of the things they seek.

3.4 Difficulties in detecting missing evidence

Investigation involves collecting evidence and forming hypotheses. Investigative abilities are a part of human nature, and investigation is an essential means for learning, e.g., from diagnosing health problems to evaluating products. However, it appears that people tend to be weak in recognizing missing evidence, and consequently, in adjusting the certainty of their hypotheses to the realities of incomplete data [Heu81].

In deception, one way to falsely portray something is to create fake evidence that implies the thing's existence. For example, a honeypot can use this technique to falsely portray a firewall and its protected subnet. The honeypot just needs to return the packets (i.e., evidence) that hackers expect when scanning such a firewall.

The bias of not recognizing missing evidence can aid the deception planner when he deceptively portrays something by creating fake evidence of it. If the planner overlooks particular types of evidence, the target may likewise overlook the omission. In the example, if the honeypot does not return all the packets that hackers' scans should receive, some hackers may simply overlook that missing evidence.

4 Impaired thinking

"The fraud specialist is expert at taking advantage of our weaknesses. He knows how to 'read' a person and assess vulnerabilities." from *The Rip Off Book : The Complete Guide to Frauds* [1]

To carry out their deceptions, con-men often exploit some form of impaired thinking in their victims. This section presents four of the forms they use: time limitations, false expectations, cravings and compulsions, and limitations in critical thinking. A fifth type of impaired thinking that is commonly exploited in physical security is also presented: a guilty conscience. The section shows how the five forms can be applied to computer-security.

4.1 Time limitations

Frauds are often "limited time" offers [San94]. Con-men create scenarios that require urgent action, so the victim does not have time to think critically about the deception or investigate it. Typically, the victim is presented with the apparent dilemma of hastily choosing now, or forever loosing the opportunity. This ploy can also be used in computer security deceptions. An example is Cliff Stoll's sting operation (see section 2.3), which involved an offer for information on nuclear weapons. The deception was strengthened by giving the offer a soon-approaching deadline.

[1] [San84]

Limited-time offers can also be used to advantage in honeypot design. When hackers discover a new network-sever vulnerability, they have a small window of time to exploit the vulnerability, as it will be promptly fixed in well-maintained networks. Hackers' haste to exploit new server vulnerabilities can make the hackers vulnerable to deception. Honeypots can take advantage of this vulnerability by using servers with recently announced vulnerabilities.

4.2 False expectations

Section 2 described how expectations influence perception and how deceptions can exploit expectations, including erroneous expectations. This section describes two specific types of false expectations that con-men exploit. Their application to computer security is also described.

First, false expectations are created whenever someone misunderstands how something works [Hus97, San84]. For example, it is not uncommon for emerging computer technology to be grossly over-valued by consumers and investors. Historical examples include artificial intelligence, Java, and even the Internet. Consumers and investors who hold such false expectations are vulnerable to deception, and the vulnerability is invariably exploited by con-men who promise to deliver the emerging technology. Similarly, many hackers will also hold popular delusions about emerging technology, and these false-expectations could be exploited to deceive them. For instance, firewalls and intrusion detection systems (IDSs) are security systems whose power and effectiveness have, historically, been widely over-estimated. It may be possible to deter hackers by using deception to exaggerate the effectiveness of a network's computer-security systems. As an example, hacking instances that are detected by conventional means could be attributed to the "new generation of powerful IDSs." Indeed, in the mid 80s, a hacker wrongly concluded that his subsequent attempts to access a computer system at SRI International failed because of an IDS. His expectations were based on reading a report about IDSs on the machine, and deducing that the concept had been implemented. In fact, the passwords on the system had been changed following his initial break-in.

Second, trust creates expectations that can make a person vulnerable to deception [San84]. If the deception target trusts something that is corrupt, or corruptible, then that trust can be used to deceive. For example, consumers tend to trust name brands, and a corporation can deceptively exploit that trust by selling substandard products under its brand. Deception occurs when the corporation allows buyers to assume the substandard product is of the same quality as its other products. Similarly, hackers rely on a variety of systems, tools and organizations, and their trust in these things can potentially be exploited. As an example, when hackers break into Unix computers, they often download and compile hacking tools. Their trust in the resident compilers can be exploited. For instance, the compilers can be rigged to create binaries that secretly trigger security alarms whenever the code is run.

4.3 A guilty conscience

King Solomon observed that "the wicked flee when no one pursues. . ." (Proverbs 28:1). Apparently, criminals have a guilty conscience, and it tends to make them paranoid about getting caught and punished. They are hypersensitive to the possibility of detection and retribution. Also, they respond fearfully. Such hypersensitivity can make them vulnerable to deceptive indicators of detection and retribution. For example, fake security cameras, and signs warning about nonexistent alarm systems, can be very effective.

In computer security, most hackers are criminals, e.g., trespassing script-kiddies, cyber thieves, and state-sponsored hackers who are engaged in unjust warfare. Hackers' guilty consciences can make them hypersensitive to deceptive indicators of detection and retribution. For instance, well publicized hacking prosecutions can be used to exaggerate intrusion response capabilities. Also, fake displays of network intrusion-detection systems can be used to exaggerate detection capabilities, as commonly done in physical security. For example, if hackers suspect honeypots are being used, real computers can be given honeypot components that hackers look for, such as a keystroke logger.

4.4 Cravings and compulsions

One of con-men's most well known techniques is to exploit greed [BW82, San84]. Greed powerfully lulls suspicion, impairs critical thinking, and thereby makes people vulnerable to deception. In general, there are a variety of cravings and compulsions that impair thinking and make humans vulnerable to deception. The causes of these cravings and compulsions include: a) moral vices, such as greed, substance abuse, uncontrolled anger, and a lust for power and fame; b) desperation, as seen by the perpetual sales of fraudulent remedies for terminal illnesses and excess weight; and c) psychological disorders, such as obsessive-compulsive behavior.

Cravings and compulsions make humans vulnerable to deception in two ways. First, they impair the thinking abilities needed for counterdeception. Secondly, when a deception offers the target what he wants, the opportunity will often arouse his suspicions. In such cases, cravings and compulsions can cause the target to take foolish risks and thereby fall for suspected deceptions.

Hackers are often characterized by their vices and disorders. As described earlier, most hackers are criminals, and consequently, they are engaged in vice. For example, many script kiddies covet the technical abilities that will make them "elite" and famous among their peers. Cyber thieves are driven by greed. Hacking itself can be highly intriguing, and hackers commonly display extreme obsessive-compulsive behavior in their hacking. A good example is the hacker Matt Singer, who was unemployed and hacked constantly [FM97].

Deception can exploit the target's impaired critical thinking, caused by cravings and compulsions. For instance, Singer's obsessive behavior seemed to impair sober-minded reflection about his vulnerabilities and risks. When his brother cautioned him about getting caught, he replied that he was telnetting through too many systems to be tracked. Apparently, it did not occur to Singer that his initial connection was often to the same university network, and its system administrator was stealthfully monitoring his world-wide hacking adventures.

4.5 Limitations in critical thinking

Another vulnerability to fraud arises from deficient critical thinking. There are two types of such thinking that con-men often exploit, and they can be used for computer security deceptions. One deficiency is *credulity*, or the willingness to believe something based on slight or uncertain evidence [San94]. A common cause of credulity is naiveté, as superficial knowledge can limit critical thinking and make one vulnerable to deception. Hackers can be quite naive about the networks they hack, due to their unfamiliarity with the network topology and the operation it supports, e.g., banking or military. Script-kiddies will tend to be credulous due to youthful naiveté. Another deficiency in critical thinking is *laziness* [San84]. It may be possible for a hacker to discover a deception, but the deception will be safe if the hacker is not willing to invest the effort required for discovery. Hackers who do not fear being caught, or who act impetuously, may simply not make the effort needed for counterdeception. Many script-kiddies are likely to act in this manner.

5 Conclusion

Table 1 summarizes the eleven psychological vulnerabilities to deception presented in the paper. Exploitation of the vulnerabilities can increase a deception’s likelihood of success. An understanding of the vulnerabilities is a tool for the deception planner’s toolbox, and the vulnerabilities’ most significant uses are recapped here. In the military and intelligence deception literature, there is a resounding admonition to exploit the target’s expectations and desires. The work of fraud artists indicates that the target’s cravings and compulsions are desires that make him particularly vulnerable to deception. In general, deceptions that are contrary to the target’s expectations should be avoided, if possible.

From our analysis of deceptions that exploit psychological vulnerabilities, we make three observations regarding their application to computer security. First, deceptions that exaggerate security capabilities such as intrusion detection can potentially exploit a guilty conscience, false expectations and all of the cognitive biases. Second, things that the target expects to be hidden can often be deceptively portrayed just by showing their indicators or evidence. Such deceptions can potentially exploit biases toward causal explanations, oversensitivity to consistency, and difficulties in detecting missing evidence. Third, deceptions based on conditioning can exploit biases toward causal explanations and biases in estimating probabilities.

There are limitations to exploiting psychological vulnerabilities to deception owing to uncertainties in the target’s reaction. Fortunately, there are several ways the deception planner can manage or reduce the problems associated with this uncertainty. First, the uncertainty can be reduced by gaining a better understanding of the targets’ psychological vulnerabilities. Second, although some psychological vulnerabilities are capricious, others are more predictable, such as hackers’ expectations about network traffic. Third, when designing deception operations, the deception planner does not have to focus on exploiting the target’s psychological vulnerabilities, but rather, he can exploit the vulnerabilities when the opportunity presents itself. Lastly, for many deceptions, the exploitation of psychological vulnerabilities does not have to work all the time, just often enough to be useful.

The savvy deception target will be familiar with psychological vulnerabilities to deception. He will seek to minimize them and to detect attempts to exploit them. For instance, his counterdeception work will benefit from the knowledge that most deceptions will seek to exploit his expectations and desires. However, to a certain extent, psychological vulnerabilities to deception are unavoidable, due to the inherent weaknesses and limitations of humans. For example, although expectations are fallible, they are a necessary means for making sense of the overwhelming information received by the senses. The target must form expectations, and these expectations can often be used to advantage in deception.

WWII deception planner Lt Col Geoffrey Barkas provides an insight into the human vulnerability to deception [Bar52]. Barkas was responsible for many of the highly successful deceptions that contributed to Rommel's defeat in North Africa in 1942. After seeing the Germans capture a dummy oil port he had built, Barkas thought the Germans would never be fooled again, as they had now seen what British deceptions could accomplish. However, further successful deceptions led Barkas to conclude that, "as long as the enemy has a good intelligence service and pays attention to what it says, it will be possible to fool him again and again." The British used the German intelligence service to communicate deception stories to the German military leaders. The Germans could be deceived repeatedly because their human limitations left them ever vulnerable to deception. In general, deception is always a possibility, as the target's counter-deception efforts cannot fully overcome his inherent vulnerabilities to deception. This often provides the deceiver with an advantage over the target. However, the advantage is not unilateral—the deceiver is also flesh and blood, and inherently vulnerable to deception himself.

Table 1 : Summary of psychological vulnerabilities to deception

Category	Vulnerability	Description	Example Exploits of the Vulnerability
Perceptual Biases	expectations	expectations filter the overwhelming information received from the senses	build honeypots that behave like real systems
	desires	goals, objectives, passions and obsessions	bait hackers into staying on a honeypot by offering documents that appeal to them
Cognitive Biases	bias toward causal explanations	tendency to see order when it does not exist	exaggerate security capabilities with deceptive indicators that are random and ambiguous
	oversensitivity to consistency	overestimation of the relevance of consistency in small samples	hide valuable computers by providing multiple forms of information consistent with them being printers
	biases in estimating probabilities (availability bias)	overestimation of things that are easily recalled or imagined; also, underestimation of things than are not as easily recalled or imagined	to deceptively portray a real system, a honeypot can portray computers that hackers can easily imagine or recall
	difficulties in detecting missing evidence	weakness in recognizing missing evidence; also, weaknesses in adjusting the certainty of hypotheses in accordance with missing evidence	when a honeypot deceptively portrays a firewall, it may be sufficient for the honeypot to impersonate most, but not all, of the packets that are returned by a real firewall
Impaired Thinking	time limitations	limited-time offers	to induce hackers to act hastily and recklessly, honeypots can have recently announced vulnerabilities that hackers expect to be promptly fixed
	false expectations	false expectations about how things work; also, misplaced trust	exploit hackers' trust in compilers by rigging them with code that triggers security alarms when used
	a guilty conscience	"the wicked flee when no one pursues"	portray deceptive indicators of attack detection and response to make the hacker "flee", i.e., to deter hacking
	cravings and compulsions	caused by moral vices, desperation, and psychological disorders	exploit a hacker's compulsions that cause him to underestimate his vulnerability to stealthy monitoring
	limitations in critical thinking	credulity or laziness	exploit a hacker's laziness to investigate whether a honeypot is real

6 Bibliography

[Bar52] Barkas, G. *The Camouflage Story*, Cassell & Co. Ltd, 1952.
[BW82] Bell, J., B. Whaley. *Cheating and Deception*. Transaction Publishers, 1982.
[CIA80] *Deception Maxims: Fact and Folklore*, Deception Research Program, Office of Research and Development, Central Intelligence Agency, 1980.
[Dew89] Dewar, M. *The Art of Deception in Warfare*, David & Charles, 1989.
[DH82a] Daniel, D., K. Herbig, editors. *Strategic Military Deception*, Pergamon Press, 1982.
[DH82b] Daniel, D., K. Herbig. "Propositions on Military Deception", in [DH82a].
[FM97] Freedman, D.H. and C.C. Mann. *At Large: The Strange Case of the World's Biggest Internet Invasion*, Simon & Schuster, 1997.
[Heu81] Heuer, R. "Cognitive Factors in Deception and Counterdeception", in [DH82a].
[Hus97] Huston, P. *Scams From The Great Beyond : How To Make Easy Money Off Of ESP, Astrology, UFOs, Crop Circles, Cattle Mutilations, Alien Abductions, Atlantis, Channeling, And Other New Age Nonsense*, Paladin Press, 1997.
[ISV95] Icove, D., K. Seger, and W. VonStorch. *Computer Crime : A Crimefighter's Handbook*, O'Reilly, 1995.
[JDD96] Joint Doctrine Division, *Joint Doctrine for Military Deception*, U.S. Joint Command, http://www.dtic.mil/doctrine, 1996.
[Jer68] Jervis, R. "Hypotheses on Misperception", *World Politics*, 20(3):454-479, April 1968.
[San84] Santoro, V. *The Rip Off Book : The Complete Guide to Frauds*, Loompanics Unlimited, 1984.
[San94] Santoro, V. *Economic Sodomy : How Modern Fraud Works and How to Protect Yourself*, Loompanics Unlimited, 1994.
[Sch93] Schlossberg, H. *Idols for Destruction : The Conflict of Christian Faith and American Culture*, Crossway Books, 1993.
[Sto89] Stoll, C. *The cuckoo's egg : tracking a spy through the maze of computer espionage*, Doubleday, 1989.
[TK71] Tversky, A., Kahneman, D. "The Belief in the Law of Small Numbers", *Psychology Bulletin*, 76:105-110, 1971.
[USA88] *FM 90-2 Battlefield Deception*, U.S. Army, 1988.
[USM89] *FM 15-6 Strategic and Operational Military Deception: U.S. Marines and the Next Twenty Years*, U.S. Marine Corps, 1989.
[Wha69] Whaley, B. *Stratagem : Deception and Surprise in War*, Center for International Studies, Cambridge, 1969.

7 Authors

Jim Yuill is a PhD candidate in the Computer Science Department at North Carolina State University. This paper is related to his dissertation. Jim previously worked at IBM in operating systems development. *jimyuill-at-pobox.com*

Fred Feer is retired from a career with the U.S. Army counterintelligence, CIA, RAND and independent consulting. Deception has been an interest and area of professional specialization for over 40 years. *ffeer-at-comcast.net*

Dr. Dorothy Denning is a Professor in the Department of Defense Analysis at the Naval Postgraduate School. She worked in the computer security field for 30 years and is author of *Information Warfare and Security*. *dedennin-at-nps.edu*