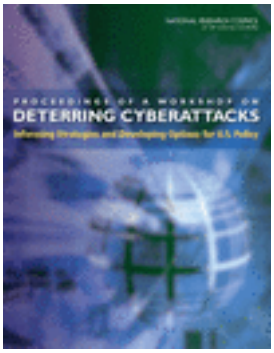

Proceedings of a Workshop on Deterring CyberAttacks: Informing Strategies and Developing Options for U.S. Policy

Committee on Deterring Cyberattacks: Informing Strategies and Developing Options; National Research Council

ISBN: 0-309-16086-3, 400 pages, 8 1/2 x 11, (2010)

Decision Making Under Uncertainty

Rose McDermott
Brown University

Decision makers are often confronted with threats and risks they may not understand, but nonetheless need to address. Among these kinds of problems, cyberattacks may present novel technical and political challenges which leaders are not adequately prepared to properly resolve. Under such circumstances, policymakers strive to reduce uncertainty, often believing, rightly or wrongly, that the quality of decisions based on more information, and more certain intelligence, will prove superior. Decisions based on more, or more certain, information may not always emerge normatively better, but decision makers may feel more confident of the choices they make under such conditions. This proclivity to favor certainty in the face of conflict exists in the realm of cyberattacks. In fact, cyberattacks constitute an attack on certainty itself; indeed, creating uncertainty in the population about the security of particular systems, such as the electrical grid, may represent part of the goal of an attack. Importantly, the central problems confronting decision makers in the realm of cyberattack remain fundamentally and intrinsically psychological in nature because they so often involve high degrees of uncertainty about the source of the attack, and the motivations of the perpetrator. Yet these psychological problems, which often incorporate technical elements about which decision makers often remain either uneducated or unaware, nonetheless require decisive political remedies. In order to achieve this goal, decision makers must learn to accept and manage uncertainty even under conditions of stress and threat.

American leaders in particular tend to look for solutions to complex problems in terms of technological fixes, since such remedies appear straightforward and direct. But such responses can fail, as often occurs in intelligence when signals intelligence is privileged over human intelligence, often to the detriment of accuracy; such a difference may explain part of the reason the Israelis were less surprised than the Americans by the fall of the Shah in 1979. Similarly, problems may exist for which no technical solution yet exists.

Several specific classes of psychological biases tend to plague decision makers confronting uncertainty. In the case of cyberattack, the tendency for aggression to become heightened in anonymous circumstances raises the risk for unnecessary confrontation, and processes of social contagion intensify the time pressure under which leaders may have to operate. Ironically, awareness of such proclivities does not necessarily diminish their occurrence. Indeed, overconfidence can further exacerbate such effects.

The following discussion outlines potential ways to reduce uncertainty, concentrating on the biases which can infuse processes of decision making under such conditions. To begin at the end, cyberdefense

and cyberoffence will remain inadequate to address the extent and scope of the impending challenge, accurate and confident attribution will prove enormously difficult if not impossible to achieve, and other technical solutions are likely to remain impossible for the foreseeable future. As a result, gaining an adequate appreciation of the ways in which systematic psychological biases surrounding attribution and confidence, in particular, can influence decision making under uncertainty in particular ways can help cut off the tendency to create the very problems we seek to ameliorate. In order to make appropriate decisions under conditions of unavoidable uncertainty, leaders should be made more aware not only of the technical nature of the challenges they face, but also of the psychological biases which may impede optimal choice when options remain limited, uncertain or truncated. Institutional strategies for reducing the influence of such biases, including efforts to increase transparency across response options, may help reduce the negative influence of these tendencies. In addition, improving expertise and calibration could reduce the prospects for disproportionate responses to cyberattack, which may lead to undesirable unintended consequences.

DECISION MAKING UNDER UNCERTAINTY

The central psychological problem in addressing uncertainty results from the fact that individuals tend to jump to conclusions prematurely in a desperate attempt to gain some sense of control, however artificial, over the situation at hand. When this happens, several consequences follow. First, leaders, like others, strive to fit a given situation into a pre-existing category they already have in mind for "what kind of conflict" this resembles, or "who is to blame" for a particular outcome. These desires, however normal, can lead to systematic biases because they encourage decision makers to fit a novel situation into a pre-existing theoretical category, which then serves to shape and structure incoming information in ways which conform to those pre-existing beliefs. For example, they readily accept information which conforms to those beliefs without sufficient interrogation, they reject intelligence which refutes those ideas without investigation and they interpret ambiguous information in ways which conforms to, and supports, those pre-existing ideas (Lord, Ross & Lepper, 1979). In other words, individuals tend to apply different criteria to incoming information based on the degree to which it aligns with their pre-existing theories of the world.

This process can be exacerbated up the chain of command when many decision makers share a particular world view, as occurs when individuals of a given generation share a powerful historical analogy, such as Munich, Vietnam, or 9/11 (Khong, 1992). These analogies then serve not only to organize information but also to provide appropriate ready-made responses, although oftentimes these alternatives do not derive from successful choices in the past, but rather emerge from what leaders thought might have solved the problem in hindsight. So, for example, those who opposed appeasement of Hitler at Munich might assume that standing firm against future potential tyrants will prevent the outbreak of war. This can work so long as the current challenge closely resembles the historical one in critically important ways; however, often the new threat contains novel elements which require greater ingenuity in generating a response. When this occurs, trying to fix a current problem with a past solution works about as well as trying to fit a computer into a typewriter case.

The second main consequence that follows from premature cognitive closure in the face of uncertainty is that the initial judgments then come to serve as anchors for future assessments of the situation (Webster & Kruglanski, 1997); under this circumstance, individuals tend to "seize" on early information and then "freeze" on those evaluations even as new contradictory information may emerge. These tendencies appear to vary across situations but represent pretty stable aspects of individual personality differences. The preference for "seizing" early seems to derive from an impetus for urgency, while the propensity toward "freezing" appears to result from a desire for permanence, or stability. When this occurs, decision makers prove much less likely to re-evaluate earlier decisions than they would if they were still making up their mind about the nature of an attack and their options for response. These initial evaluations, often based on very little information but tremendous urgency to achieve certainty,

will be heavily influenced by prior beliefs and images regarding the nature and capabilities of given adversaries.

These tendencies will likely exert themselves strongly when confronting the prospect of cyberattack because the nature of the threat is reasonably new and many leaders may not be familiar with the technology or its level of complexity. Knowing how to prepare people to respond without necessarily having to teach them the technical aspects of the problem will prove challenging. Leaders may be unsure how much information they need in order to be confident regarding how best to response. They may have very clear suspicions about who constitutes an enemy, and what kinds of responses will deter which kinds of attacks, but they may not have enough information to go public with their claims.

ENSURING APPROPRIATE ANALOGIES TO GUIDE DECISION MAKING

The goal of preventing cyberattacks against the United States is important, but the presumptive analogy of deterrence upon which it rests seems misguided, if pervasive. Deterrence as a theory developed in the context of kinetic attacks which allowed some attribution, albeit often biased, regarding instigation and intent. But unlike most military attacks, cyberattacks defy easy assessments of perpetrator and purpose. Deterrence dominated as the strategy of choice during the Cold War between the United States and the Soviet Union. However, cyberattack looks much more like the current threats posed by non-state terrorists than like those posed by hegemonic state actors. As with terrorism, the very point of cyberattack, at least in part, is to increase uncertainty. Applying possibly anachronistic notions of deterrence to such a threat may lead decision makers away from the most creative approaches, and constructive responses, to these challenges.

One of the main problems with deterrence was noted by Robert Jervis (1979, 1982) in his work on nuclear deterrence; the fundamental logic of deterrence rests on assumptions of rationality which do not accurately depict human decision making processes. Deterrence depends on a credible threat to respond and deny the adversary any advantage from an attack, and yet such an ability requires decision makers to clearly give and receive clear signals of intent and capability. In the context of nuclear weapons, Jervis noted the many psychological obstacles that might prevent leaders from properly evaluating the opponent's values, credibility, or perceptions; all these calculations provide the basis upon which reliable deterrence depends. Recent evidence from psychology and neuroscience only serve to validate and clarify such reservations regarding human decision making.

The misuse of the deterrence analogy should prove instructive to those who remain skeptical that leaders fall prey to such inappropriate historical references. In the case of cyberattack, a more apt analogy might be drawn from the risks posed by pandemic disease and the principles of disease surveillance employed to try to contain such threats. Pathogens co-evolve in the human immune system, often operating as parasites, but also working to facilitate useful functions, such as digestion. From a public health perspective, many pathogens only constitute a threat when they become unusually virulent or prevalent. Similarly, most people can use their computers normally without it posing a risk to the broader community. However, when a particular computer gets taken over, especially without its owner's knowledge, the person who controls the machine can wreck public havoc, just as an individual can spread disease prior to actively experiencing symptoms himself. While epidemiologists need not broadcast their surveillance of disease prevalence, and such watchdogs rely on primary care physicians for providing first responder information, systematic monitoring remains crucial to any attempt to keep potential epidemics under control, or to limit their spread and destructiveness once they emerge. In the case of a pandemic disease such as the flu, public health officials try not to kill infected individuals, but rather seek to kill the dangerous pathogen. Efforts to monitor and control the emergence of botnets can serve a similar goal of providing firewall protection to the wider public by containing the dangerous computers, rather than going after those who control them remotely since these individuals are much more difficult to identify and locate.

The specter of cyberattack raises similar concerns regarding attitudes toward regulation and surveillance as well as the nature of underlying security relationships. The influence and effect of some cyberattacks, like some diseases, may only become fully evident over time. Some attacks may initially appear worse than they are, like Legionnaire's disease, while the lethality of others, such as AIDS, may simply prove incomprehensible upon first appearance. Like a lethal pathogen, the adversary in cyberattacks is often supremely adaptive and quick to respond to efforts to stop their grown and activity. Hardening defenses often only serves to redirect the energy of the assailant to more vulnerable areas. As with disease, in the case of cyberattack, it may prove less effective and efficient to go after the adversary than to simply target the infected machines, thus denying the adversary the means of delivery.

APPROACHES TO REDUCING UNCERTAINTY

When confronted with uncertainty, leaders will naturally seek to reduce that uncertainty by obtaining more information about the nature, source, intent and scope of the perceived threat or attack. There is nothing wrong with this inclination; problems arise when such intelligence is treated with a degree of confidence which the evidence cannot support. When subsequent actions are then based on these inferences, numerous negative consequences can arise, from increasing the risk of blowback by potentially harming innocent parties, to initiating actions that end up harming one's own systems, to escalating a conflict beyond expectation or desire. Awareness of the processes by which such decisions might become skewed can help to design institutional constraints and checks, as well as foster greater transparency and accountability.

When responding to a potential cyberattack, leaders confront a variety of questions, not least those relating to the source and intent of the attack. What is the purpose of the attack? Who instigated it? Can we stop it? Should we retaliate, and, if so, how? In striving to answer these questions, leaders have a variety of tools and strategies available to them; however, basing such decisions on an atavistic model of deterrence only serves to restrict creative responses to these challenges by enforcing an analogy based on assumptions of kinetic systems and psychological rationality which do not apply in the cyber realm.

There are many ways in which these assumptions can undermine optimal decision making, defined as maximizing democratic goals of peace, prosperity, freedom and openness in the context of working actively to suppress any given threat. There are at least three domains where leaders can, and should, strive to operate effectively to reduce the risk posed by those who would try to disrupt American cyber capabilities; each of these areas can remain plagued by uncertainty. The first relates to issues of offense and defense, the second to questions surrounding attribution, and the third revolves around additional technical solutions.

One way in which leaders try to reduce threat is through the use of offensive and defensive counter strategies, which can be active or passive in nature. In current discussions, defense tends to solicit most of the attention, with a lot of effort going into passive efforts, such as establishing anti-virus software, firewall protection, or improving password strength and encryption. However, these strategies can often create the very problems they are intended to prevent, as occurred with the recent McAfee security update error which precipitated a global crash of tens of thousands of PCs.[1] Active defense work, which in the case of cyberattack may involve some kind of offensive counter-attack, but could also include more typical police investigation strategies, is also available to leaders. Offensive strategies, including cyberattack and cyberexploitation, on the other hand, tend to be classified in nature.

Yet clearly attributing the identity of an attacker in a cyberattack scenario presents one of the most difficult and salient challenges facing defenders. This issue of identification and responsibility often lies at the crux of the uncertainty facing decision makers in determining the nature of the threat they confront, and the possible responses they might take in reaction. Note that such problems pose threats not only for the United States, but also for other countries which might react to attacks as well. Foreign

[1] http://www.wired.com/threatlevel/2010/04/mcafeebungle/.

leaders may also share similar constraints and uncertainties in their decision making as well, and U.S. leaders should be careful not to assume that adversaries or allies have a clearer understanding of attribution, or how best to respond to it, than we do. How such attacks affect others can also compound the effect on U.S. systems; there will be inevitable interconnections in how their responses affect our own systems, and U.S. defenders may need to stay attuned to such interaction effects.

The problem with attribution in cyberattack, of course, is that even if it is possible to locate the IP address of the attacking machine, it may be impossible to know who exactly is controlling that machine. Even if investigators find a particular location, they may never be able to locate the computer which launched an attack, as for example might occur if a Starbucks with open wireless networks served as the instigating location. Plausible deniability results from the very real possibility that a given computer was hacked and taken over by another with no knowledge on the part of the host about the nefarious activities his machine subserved. Because the knowledge required to learn how to become an expert hacker, unlike that required to make a nuclear bomb, for example, is relatively accessible and inexpensive, it may be impossible to determine the source of any given attack. Clever attackers can use many machines as stepping stones and hide their electronic tracks with still other machines, and victims may remain quite limited in what they can do to overcome it, undermine it, or learn about its original source (Stoll, 2005).

Efforts to improve and strengthen authentication can only take investigators so far because assertions of identity requires some kind of credentialing that may itself seem untrustworthy. If, for example, the Google attacks from China were launched from various students at two different universities, would the United States be more likely to believe they had nothing to do with the Chinese government if they had credentials from the Chinese government saying they were merely students and not government operatives? Just as forensic computer programmers may not trust any information they obtain from a machine they judge to have been compromised, and taken over by another, agents of the United States government may not trust any government-issued identification as either authentic or accurate. As a result, it is unclear how much effort is worth dedicating to processes of authentication. Such validation only works in a larger context of trust, whose very existence would vitiate the need for such credentialing in the first place. In order to understand this phenomenon intuitively, it may help to consider the mirror counter-factual experience. Imagine, for example, how the United States might expect other governments to respond to a claim that a hacker was working entirely independently out of an American city. Would the government expect Britain to believe such an argument more than, say, China?

Attributions of identity matter because they are used, at least in part, to infer intent. The United States may not like it if it finds the British or Israelis penetrating our defense IT network, but the sense of threat and fear will be much less than if similar acts were attributed to the Chinese or the Iranians. And yet an attack emanating from Iraq might appear to come from a computer in London just as easily as from one in Tehran. In one of the first coordinated cyberattacks discovered in early 1998, although apparently going on for a year or more prior, the so-called Moonlight Maze incident, an attack on the United States Department of Defense, NASA, and other government agencies appeared to emanate from Moscow, although it remained unclear whether that was the original source of the attack. The Soviets denied any involvement. But the identity of an attacker remains central to making inferences about intent, just as the apparent goal of an attack can provide information about its potential source. This inferential process highlights the importance of the influence of emotion on assessments of uncertainty and outcome; fear and anger affect predictions of probability in systematic ways which will be addressed in greater detail below. Suffice it to note here that fear leads to more pessimistic evaluations, while anger produces more optimistic assessments regarding future outcomes (Lerner & Keltner, 2000).

Another reason accurately identifying an attacker can prove critical is because any American president, or other leader, may remain reluctant to take a decisive action in retaliation for an attack if the source of such an attack remains unknown or uncertain. Yet accurate assessments take time, providing another reason to diminish the basis for effective models of deterrence to function. If the source of an

attack cannot be identified with certainty, how can a counterattack, either a cyberattack or a kinetic one, be initiated without risking the possibility of unwanted collateral damage or potential blowback?

But sometimes leaders may feel more confident of their attributions of villainy than the objective evidence might support, particularly if they have jumped to conclusions based on judgments resulting from the salience of recent attack. In this way, leaders may emphasize the utility or destructive consequences of a potential attack over its probability. By highlighting potential damage in making risk assessments, and privileging value over probability in making decisions about how to respond, decision makers may adopt a worst case scenario approach to a cyberattack and jump to unwarranted conclusions out of an understandable desire to prevent future damage. If a leader gets this assessment wrong, his estimate of the possibility that a particular perpetrator instigated the attack will be higher than if his assessment was correct. The interaction between these biases can thus complicate accurate judgments concerning attribution.

However, technical analysis alone is unlikely to allow a decision maker to accurately speculate about the source of a cyberattack. Intelligence will most likely fail, and even sophisticated observers may never know if what they see represents the actual source, or whether someone else has engaged in deliberate deception to make it appear that a proxy machine constitutes the real source of an attack. Indeed, one of the sequelae of premature cognitive closure in the face of threat is that leaders will be more prone to assume threats will emerge from places where they have seen similar attacks emanate previously. If a secondary attack closely followed the one launched against Google, assumed to originate from China, it would be much easier for decision makers to assume that the subsequent attack similarly emerged from China. For this reason, catalytic conflict, or conflict between two countries initiated by a third party, remains a significant concern. Yet such attribution may or may not prove accurate. Uncertainty regarding source, and often intent as well, remains inherent in the domain of cyberspace itself, inextricably bound to the nature of the technology itself.

Moreover, lest observers appear naïve, it may remain in the best interests of the United States government to be able to hide behind attributional equivocation and uncertainty, not only in order to engage in unacknowledged retribution against attackers, but also to protect concealed sources whose skills may allow the United States to conduct powerful offensive cyberattacks as well. Such ambiguity might serve at least two important purposes. One, it allows the United States to engage in the kind of plausible deniability about initiating attacks which, while criticizing other governments for so doing, nonetheless allows the government to take aggressive action against potential threats while keeping the means of such attack secret. Second, a presumed veil of ignorance surrounding the source of attacks might allow the U.S. government to take back door advantage of presumed aggressors, or even try to take over particular machines, much like intelligence agents tried to turn spies into double agents during past wars.

Yet, as with the brief superiority which the United States had with regard to Multiple Independently Targetable Re-entry Vehicles (MIRV) technology in the days of the land based Intercontinental Ballistic Missile arms races with the Soviet Union, dominance may eventually give way to parity, as is likely already the case in this arena. Or, in the case of cyberattack, the most sophisticated technology could also reach into the hands of terrorists who may not be content to only try to steal money or information, but rather seek to wreak mayhem and destruction on the power grid, air traffic control, banking, and military command and control. As with the use of any other mercenary, states and other actors can gain rapid advances by paying skilled individuals to sell what they know to the highest bidder. Such prior experiences in the kinetic domain encourage caution and skepticism about the value of seeking to achieve or retain dominance in any rapidly changing and progressing technology.

PSYCHOLOGICAL ISSUES IN ADDRESSING UNCERTAINTY

The fundamental problems related to decision making under conditions of uncertainty, such as those which remain pervasive in the nature of cyberattack, involve psychological issues. And the appropri-

ate remedies need to rest on primarily political, as opposed to technical, grounds because the inherent risks affect political values and calculations, and often remain immune to any quick technological fixes designed to assure attribution. There are several significant psychological insights which can be brought to bear in an analysis of this topic. The most relevant include biases surrounding judgment under uncertainty, attribution, contagion and social validation, and confidence. Each will be discussed in turn below.

Judgment Under Uncertainty

In the original psychological work on judgment under uncertainty, Slovic, Kahneman & Tversky (1982) described three biases which they found systematically affect the way people render assessments concerning future probabilities. To be clear, in combination with their work on decision making, most prominently explicated under the rubric of Prospect Theory (Kahneman & Tversky, 1979, 1984; Tversky & Kahneman, 1992), they made a distinction between judgment under uncertainty and decision making under conditions of risk. Judgment refers to those assessments about the likelihood that a certain event or person belongs in a particular category, or probability of a particular event occurring in the future, for example. These judgments primarily constitute assessments of an objective reality about the real external world. Decision making, on the other hand, refers to the primary internal calculations that individuals make, which can involve trade-offs between values, as the classic guns versus butter debate best exemplifies in the political realm. Decision making under risk occurs when the decision maker faces a threat to something of value; the actor must contemplate how much of something of value might be lost if different alternatives are chosen. The right choice may allow the person to move forward in a positive way, but the wrong choice may precipitate the loss of important values, including money, status, reputation, or some other significant thing of subjective worth or import. Judgmental biases relate to the part of a rational model that refers to probability; decision making biases mirror the aspects of a rational model that refers to assessments of utility. This distinction remains meaningful in analytic terms when describing the relevant judgmental heuristics which can bias judgment.

The original work on judgmental biases, which remains robust in its demonstration of these processes, identified three primary heuristics, or basic rules of thumb individuals typically utilize in rendering predictions about the likelihood or frequency of events. By and large, these biases work effectively and efficiently, particularly in the midst of the ecological context in which they evolved, but they can lead to systematic and predictable biases, particularly when abstracted to unfamiliar contexts (Gigerenzer, 1996).

The first, representativeness, showed how people rely on similarity to make judgments about whether a particular event or person fits into a specific category; profiling individuals to determine whether they are likely to be terrorists represents a systematic attempt which depends, at least in part, on this same kind of procedure. The second, availability, suggests that individuals rely on ease of accessibility and imagination to judge how likely a future event might be. Attempts to protect the airlines from a suicide terrorist represent an example of the ways in which salient experiences not only bias judgments of probability, but also restrict imagination of future events to overweight those which have occurred in the past. The last, anchoring and adjustment, refers to the way in which people tend to gravitate to a particular focus or target, even when they know that it is irrelevant to the judgment they are being asked to make. They then often fail to make proper accommodations for the ways which the event or individual they are being asked to judge differs from the original target. This proclivity provides a basis for the use, and misuse, of historical analogies to explain and analyze modern crises. Similarly, this dynamic explains, in part, why leaders often try to fight a current war using the strategies they believe would have succeeded in the last one.

Two specific so-called fallacies merit particular discussion in light of the issues raised by cyberattack. First, the conjunction fallacy has been shown to operate in many contexts, including political forecasting, estimates of criminal activity and decision making under risk (Tversky & Kahneman, 1983). In basic

probability theory, no combination of factors can be more likely than either of its separate components. Yet both representativeness and availability can operate to make such a combination appear more likely than either of its parts. As a result, people may judge particular events or scenarios more likely than they are in reality either because they appear similar to something familiar (i.e., this is what happened in the Estonian attacks), or because they are salient for some reason (i.e, Google just sustained a similar attack). When this happens, observers can overestimate the probability of a particular event that seems more representative and available and simultaneously underestimate the probability of events which appear less representative and available. This phenomenon exerts its pull even when true probability is uncertain or even unknowable.

This dynamic becomes relevant not only in thinking about how leaders make judgments about whether a particular person or country, like China, belongs in a particular category, such as enemy or hacker, but also in rendering assessments regarding the likelihood that a particular kind of event, like a computer virus, results from either relatively benign or malicious intent. And even if malicious intent appears clear, the category of intrusion can matter in making decisions about how to respond: Is the intruder after money, information, or seeking to cause destruction? Each of these kinds of assessments requires placing individuals or actions within particular categories whose base rate probability remains unknown or unclear. Under such circumstances, similarity and salience can drastically affect assessments of probability in ways which do not accurately reflect actual likelihood, but can nonetheless profoundly influence political decisions about how to respond to a given threat or attack.

A related dynamic, called the disjunctive fallacy (Bar-Hillel & Neter, 1993), occurs when individual rank the probability of a category more likely than its superordinate categories. For example, when someone says they think it was more likely they saw a dog than an animal, or more likely someone is from Brazil than South America, they fall prey to his fallacy. These assessments, as with the conjunction fallacy, fly in the face of standard probability theory, but possess a strong psychological pull. In each instance, estimates of probability can be drastically affected in systematic ways by psychological processes which do not rest within the realm of acceptable normative models of rationality or probability. In such a way, a university source of attack may be judged less likely than a state sponsored attack, although the actual ability to distinguish between such sources remains more apparent than real.

In addition, these processes can interact with emotion in ways which exacerbate these biases under crisis situations involving fear or anger. For example, angry people tend to have more optimistic assessments of future probabilities (Lerner & Keltner, 2000); if the attacks on 9/11 made a person very angry, he would have been more likely to believe that responding with violent action would prove successful in stopping the terrorists than someone who was made sad by the same event. On the other hand, fearful people evince a more negative assessment of the future, as might be expected by those who worry about the worst coming to pass. Similarly, in considering potential responses, emotions can significantly affect a leader's decision regarding the nature of appropriate response (Sabini & Gault, 2000). Angry individuals appear much more likely to seek retributive justice, while fearful ones tend to support more rehabilitative types of outcomes. Thus, a leader, and a public, more scared about future attacks will likely support quite different choices than those who, in contrast, remain angry.

Attribution

One of the main problems with attribution concerns assessments of agency. Computers represent disembodied extensions of the actors who control them; these actors, in turn, can disavow any knowledge of any given interface, thus ensuring plausible deniability regarding both action and intention if victims manage to trace a given source.

While it may not seem entirely viable to work on developing various procedures which might allow for a higher probability of accurate attribution in such attacks, research on attribution can speak to the way in which technical and non-technical information can be combined to triangulate in on greater

confidence surrounding the identity of an attacker. The analogy here is similar to that of epidemiology, using patterns of infection to trace a disease back to its source.

Attribution theory (Kelley, 1967) focused particularly on consensus, distinctiveness and consistency as the three key feature of appropriate causal inference in attribution. While this work did not specifically seek to discern identity, the lessons apply to such attributions as well as those regarding other causal inferences. Consensus emerges when many observers make the same judgment concerning causality. Distinctiveness occurs when an actor engages in a behavior that diverges from what other actors in that same situation do, thus making their actions unique or distinct in some way. Consistency refers to the notion that a given actor does something that appears in keeping with behavior she has undertaken in the past, or seems in line with basic tendencies that observers have witnessed from that person before, or in other situations. A combination of distinctiveness and consistency may look like what a poker player would consider a "tell" in an opponent, an idiosyncratic and reliable indicator of intent or, in this case, identity.

In trying to apply these theoretical notions to attributing identity in the case of cyberattack, observers might use these same characteristics to ascertain the identity of an attacker by employing a combination of these categories. It is important to keep in mind, for example, that actors can identify online in a variety of ways. They can visually self-identity (i.e, with a picture), they can verbally self identify (i.e. with a name) or they can discursively self-identity, as when they provide identifying information, or a verbal tell, which can uniquely be associated with a particular actor, or type of actor. Such revelations may be as simple as always mis-spelling a particular word in the same specific way. The first two might be faked, but the last may be unconscious, or, at the very least, put forward by someone who has close familiarity with the target actor.

In trying to triangulate on the identity of a particular perpetrator, using attribution theory for a guide, and perhaps focusing on more discursive forms of self-identification, observers might examine the extent to which an attack is consistent with prior attacks from known sources, the degree to which allies and other trusted sources with access to the same information converge on the same putative source, and the manner in which the attack appears distinctive to a particular actor, or embodies a specific Modus Operandi for a known actor. Such a process can help hone in on a particular locale, organization or actor. This is similar to using background information in an audio or video tape to try to identify the location of the speaker.

Other Social Psychological Processes: Anonymity, Social Validation, and Contagion Effects

In their astute analysis of some of the social psychological effects which fed the First Internet War, which involved a distributed denial of service attack on Estonian government sources for several days in early 2007, Guadagno, Cialdini, and Evron (2010) pointed to several significant social psychological factors which, while not unique to cyberspace, certainly have effects which can be magnified within it. The endemic character of these forces retain prominence in circumstances of cyberattack to a greater degree than in kinetic warfare primarily because of the speed and vast potential dispersion of Internet effects. When something goes viral in cyberspace, millions of people can receive information almost instantaneously. This level of simultaneous and instantaneous exposure to mass experiences remains essentially unprecedented in prior human history. Furthermore, an attack which degrades communication may itself render an effective response to such an assault more difficult to achieve.

Anonymity provides a key to both the aggressiveness that can manifest in cyberthreats, as well as the difficulty of tracing attribution. Those who interact online have more anonymity than those who see each other face to face precisely because such visual cues are absent (Bargh & McKenna, 2004); indeed, many individuals experience a sense of depersonalization when interacting with people who are not there (Postmes et al., 2002), which can help explain why increasing time spent online can increase anxiety, diminish social capital, and heighten experiences of alienation. Yet, interestingly, Derks et al. (2008)

found *greater* frequency and explicitness in emotional displays in computer mediated communication than in face to face interactions. This may partly result from the need to convey emotions explicitly online which would be evident through non-verbal expressions and vocal intonation in personal interactions. While such emotional mediation may help explain why more than 3% of last year's marriages began with eharmony, it also reflects the fact that hostility, as well as attachment, is more frequently and easily expressed in anonymous and mediated contexts as well.

Social psychologists have suggested that anonymous forms of interaction, such as those that rule cyberspace, decrease individuals' attention to their normal standards of behavior, releasing inhibitions which might govern their behavior in face to face interactions (Joison, 2001; Matheson & Zanna, 1989; Sassenberg et al., 2005).

These models rest on earlier work on the effect of deindividuation on the propensity for violence. Deindividuation, such as occurs in large groups, or when anonymous engagement is possible, can increase the likelihood that individuals will disregard their own internal standards and espouse the established group social norm, even if that norm encourages aggressive actions. In the case discussed by Guadagno et al. (2010) in Estonia, they cite this process as one of the reasons inciting Russian sympathizers to join in a dedicated denial of service against Estonian government agencies. The medium may be different, but the process is identical to that employed by members of ACT-UP in the late 1990s with their strategy to phone-bank the FDA, denying access and egress, until the agency agreed to allow AIDS patients to try experimental drug therapies on a fast track approval basis (Carpenter, 2002).

The social identity of deindividuation model (SIDE) (Reicher et al., 1995) has been investigated in the context of computer mediated interaction to examine the effect of minority influence on majority opinion in such contexts (Moral-Toranzo et al., 2005). SIDE argues that under anonymous conditions, individuals become more attentive to the group than to themselves, and thus more likely to behave in ways which both manifest and represent group norms. This model suggests that anonymity will potentiate minority influence under conditions where group identity becomes salient, but these predictions were not sustained in the empirical work. This may be at least partly because the majority opinion represents the modal group norm to which members predominantly adhere.

Furthermore, the rapidity and broad reach of the Internet potentiates and speeds the influence of social validation and contagion effects. Social validation effects grow out of an adherence to the group norms which dominate and persevere in deindividuated environments such as those which rule the Internet. Along with liking, social proof provides one of the most powerful tools of persuasion available (Cialdini, 2001). Together, activities which make others like us for doing something everyone else is doing become particularly attractive. People are much more likely to do something once they see others engaging in the same behavior; this phenomena helps explain a lot about the cascades that precipitously occur in all sorts of social processes from economic markets to morale in combat (Rosen, 2005) where one person's behavior, or that of a small group, quickly grows into a contagion effect. As when videos of great interest go viral, so too can acts of mayhem and destruction, as long as careful instructions and sufficient motivation are provided by sophisticated organizers.

Social observers (MacKay, 1841) going back to Freud (1922) sought to understand the nature of mass hysteria, panics, and contagion effects, each of which represent a kind of social proof and validation. Interestingly, Heath & Heath (2007) demonstrate that some ideas are much more likely to prove contagious; those which contain an emotional element, such as disgust, appear to possess the ability to both spread more rapidly and endure longer in popular culture than other constructs. But clearly Internet technology potentiates the speed and rapidity with which any idea which hits the zeitgeist can spread.

The speed and breadth of these effects influence many social processes, but by far one of the most important relates to recruitment efforts. The social manipulation of identity becomes infinitely easier when those who orchestrate such identities do not necessarily have to conform to them in person in a physical way. Hitler was remarkable for his ability to perpetuate an Aryan dream when he himself was short and dark; few people have the charisma, courage or craziness to try to pull off such discrepancy

between idea and image on a large scale. But such a maneuver is much easier to pull off when the discrepancy remains out of sight. A tactical leader can recruit potential suicide bombers using extreme religious rhetoric for strategic political reasons, and do so while sitting in a fancy house filled with materialist possessions, just as easily as an old, fat, overworked grandmother can change diapers while working at a job doing phone sex; in such a context, "oh, baby" takes on an entirely different meaning. Nature abhors a vacuum and fantasy can fill in anonymous spaces with the beliefs individuals want to attribute to one other. Sophisticated actors can generate social movements to engage in dedicated denial of service attacks by hitting upon emotional appeals that motivate myriad respondents to take action; contagion effects then operate to recruit enormous amounts of labor instantaneously.

One of the most important aspects of these social psychological phenomena and processes, including those related to attribution, is that targets and observers are unlikely to know about, become aware of, or understand these processes or how the work. This means, among other things, that both groups are much more likely to assume that large states or other main actors are responsible for large attacks than may be the case, simply by the natural tendency to associate a large effect with a big causal force or actor. This belief itself represents a form of representativeness, assuming, by the logic of similarity, that big effects must spring from equally big and powerful sources, when this may not necessarily be the case. Particularly in the realm of cyberspace, where large effects can be perpetuated by a lone individual at low cost, such assumptions can prove dangerous to veracity.

Confidence

Confidence can influence the response to attacks in two ways, either by underconfidence or overconfidence. The latter is by far the more prevalent phenomena, but both effects operate in distinct ways analytically. Underconfidence can afflict a leader who tries to attribute the source of an attack to a particular entity. As with the police who may know who committed a crime, but not be able to prove it, underconfidence can afflict a leader who believes he knows who is responsible for a cyberattack, but who does not have enough evidence to be able to publicly justify a hostile response should something go wrong.

Underconfidence can also precipitate the emergence of worst-case scenarios, as when an intrusion is uncovered the intent of which remains unclear. Underconfidence may encourage a leader to assume that he must treat the invasion as an attack, and not, for example, as an attempted exploitation, because the *political* effects of appearing to do nothing may be untenable. If that is the case, the tendency to react as though every intrusion represented a worst case scenario may lead to over-reaction in certain cases.

However, the experience of overconfidence appears much more endemic and destructive in nature. Overconfidence refers to the belief that you are better and more skilled than you actually are in reality. In work on aggression in simulated war games Johnson et al. (2006) found that overconfident people held much higher expectations about their likelihood of success in combat and thus proved more likely to attack, even without prior provocation.

Significantly, this study found tremendous overconfidence among male participants in particular, who remained remarkably resistant to appropriate recalibration in the face of failure. Subjects were asked to estimate how well they would do in the game prior to play. Most men claimed that they would do extremely well, while most women estimated that they would perform at about average level. Significantly, once participants had information about whether or not they won or lost their particular game, calibration based on feedback did not fall in line with normative expectations. Women who lost recalibrated strongly downward, but did not make sufficient adjustment upward in the wake of victory. Men, on the other hand, tended to recalibrate quickly to winning, and in fact tended to overestimate their newly demonstrated skill, while failing to downregulate adequately in the face of failure.

These findings remain largely consistent with the overall literature on overconfidence and conflict (Johnson, 2004). In their evolutionary model of overconfidence, Johnson & Fowler (2009) show how overconfident individuals will make more challenges in uncertain environment. They also illustrate how

environments with lots of resources can thus prove more likely to create more conflict. They suggest that overconfidence makes evolutionary sense by maximizing fitness whenever the value of resources exceeds twice the cost of competition. Of course part of the challenge posed by uncertainty is that participants often do not know ahead of time the exact value of the resources at stake, or the cost it might extract in order to protect them. This challenge of such a calculus represents precisely why uncertainty can impose such a high psychological cost on decision makers seeking to protect central values at the lowest possible cost without incurring undue waste.

Ironically, overconfidence often tends to go hand in hand with incompetence (Ehrlinger et al., 2008). People tend to grossly overestimate their skills in all kinds of social and intellectual tasks. This appears to result from the fact that they lack the skills necessary to realize their own limitations. Failure to see the nature of their own mistakes leads incompetent people to have higher estimates of their own performance than is warranted by objective judgments. Thus, incompetence exacerbates overconfidence to the detriment of well-calibrated decision making.

Uncertainty of Outcomes

One of the real psychological challenges in trying to understand and appropriately respond to a cyberthreat derives from the kind of uncertainty related to potential outcomes. A National Research Council report (2009) outlines several types of uncertainties in this area: those that result from the use of any kind of weapon; those that derive from the use of a new weapon (what Slovic (1999) might refer to as "dread" risk, related to fear of the unknown); and those that emerge from unexpected interaction effects between military and civilian communication systems.

The first kind of uncertainty raises the specter of collateral damage, which in the cybersphere can be quite large. Businesses, hospitals, utilities and many other aspects of the commercial sector now largely depend on a functioning Internet in order to conduct their daily affairs; any action taken to deny access to an attacker could result in serious damage to interests the government seeks to protect. Also important in a democratic, free-market society is that potential technical responses that might provide greater potential protection in some circumstances, may also pose greater threats to both privacy and freedom. In addition, these processes might also affect the profitability of large sectors of the economy, such as the Starbucks mentioned above, which rely on open and inexpensive access to the Internet, not least for purposes of shopping.

The second kind of uncertainty poses the risk of blowback, whereby in an attempt to block or attack an intruder, unexpected consequences can emerge which come back to harm or destroy one's own machines, possibly even damaging the possibility for such machines to locate and desist future activities on the part of the invader. Much like the use of a biological weapon in war, which can risk damage to one's own troops depending on the way the wind is blowing, the opportunity to uncover the existence of many possibly unknown links between intermediary machines may outweigh the expected benefits of denying access, especially if such benefits cannot guarantee permanent destruction of the enemy.

The third type of uncertainty clearly overlaps with the first, but also raises additional concerns related to unpredictable system effects (Jervis, 1998), whereby unintended consequences can emerge from the attempt to separate intrinsically intertwined systems into components whose manipulation will not undermine a highly interdependent system.

CONCLUSIONS

Clearly, the risk of cyberattack, like other forms of terrorism, poses an important challenge for future decision makers. The traditional notion, built around kinetic weapons, of using deterrence to prevent an attack may not prove a viable strategic model because it rests on assumptions of rationality that do not reflect accurate notions of human decision making capability. For deterrence to work, actors must

believe that they can credibly threaten retaliation that would deny the attacker the spoils of victory. Yet such a notion rests on the belief that leaders can both identify attackers and respond in kind. Neither of those assumptions credibly holds true in the world of cyberspace. A more appropriate analogy rests on a public health model of disease prevention, where defenders engage in active surveillance to discover and monitor infections early, thus reducing the risk for lethal catastrophe.

Technical solutions may improve the ability to attribute the source of an attack to a particular machine, but there is no technical way to credibly link any given machine to a particular actor who can then be held accountable for his actions. That plausible deniability remains inherent in the anonymous nature of the technological systems themselves. Further, defensive technologies fail as offensive strategies overcome them, just as has occurred over time with kinetic weapons. While the United States and other countries such as China clearly seek dominance in cyberattack capabilities, ultimately such dominance, and the ability to maintain it, appears ephemeral in nature since so many people can acquire the skills to hack into machines and use other machines as proxies, with resources they can secure on their own at relatively little expense. The cost and sophistication of a machine does not remain commensurate with the obstacles that confront those who seek plutonium for purposes of building a nuclear bomb. Such a reality also increases the difficulty of finding and thwarting the many actors operating in this realm for different purposes. Therefore, if offensive and defensive technologies will ultimately prove inadequate, and adequate means of attribution remain impossible, decision makers are left in the realm of having to respond to attacks which by their very nature impose, and are intended to incur, high degrees of uncertainty.

If such is the case, then the game that is being played between attacker and target is primarily psychological in nature, even if the solutions need to be political in response. Under conditions of uncertainty, several psychological processes exist which exacerbate the decisional challenges which confront leaders under even the best of certain conditions. First, judgmental biases work to make certain scenarios appear more plausible than probable because of their similarity to known or previous events, or because of their salience. Emotions such as fear or anger can exacerbate the desire to respond in a conciliatory or hostile manner, respectively. Second, attributional strategies related to assessments of consensus, distinctiveness and consistency can be used to try to get at least an internal sense of likely, if not definitive, attribution, based on discursive identity revelation. Third, many social psychological processes, including anonymity, deindividuation, social validation and contagion effects, work to enhance the prospects for terrorist entrepreneurs to recruit followers to engage in destructive processes, including dedicated denial of service attacks. Careful thought should go into the best way to mobilize similar public forces in service of the national interest; greater public awareness of the stakes in simple, emotional language (i.e. "How long could you live without email?") might serve to balance the forces on this front. Fourth, overconfidence poses a real threat to obtaining accurate estimates of the likelihood of success in counterattack, and increases the probability of undertaking hostile retaliatory action without adequate or accurate information. Such responses may precipitate the kind of collateral damage and blowback which may serve to damage the very elements of society that the government is supposed to protect.

Importantly, all these forces combine to produce an aggregated effect whereby leaders will tend to leap to premature conclusions when faced with the threat or reality of cyberattack. This will happen particularly under conditions of stress, or when leaders fall prey to either groupthink, or its opposite but equally destructive force, emotional isolation (Bar-Joseph & McDermott, 2007). They may quickly jump beyond the evidence available, often with a great deal of confidence in the accuracy of their judgments, to reach conclusions about who is responsible and what they are trying to accomplish. Such beliefs will then prove unusually resistant to modification in the face of later information or further analysis. Furthermore, such ideas will be heavily influenced by the beliefs and images such leaders hold about the potential strength and malignant intent of various enemies. In this way, context will provide critical information by establishing the category into which a particular threat or attack is placed. However, such assessments may also convey inaccurate indications regarding the nature of a current attack by highlighting the source of a recent similar threat, or stressing the influence of a particular actor. These

judgments can thus skew the perception of leaders who strive to achieve certainty in order to know how best to respond. Appropriate responses may not require certainty to be effective, but leaders are likely to be reluctant to make decisions without such assurance, however false, inflated or self-created.

Given what we know about the psychological propensities to which individuals appear prone, what mistakes are leaders most susceptible to making in the area of cyberattack, as opposed to another realm? The two most significant relate to anonymity, and its tendency to increase aggression through processes of deindividuation, and social contagion, because of the speed and extent of spread which can occur on the Internet. Importantly, the effects of anonymity do not only confer to the aggressor, but also the target. If a leader has no personal knowledge of his opponent, it may be easier to order more extreme responses than if he was aware of mitigating or explanatory circumstances on the part of the assailant. Similarly, social contagion can exert effects far beyond the ability of any government to control, since the diffusion can take on a life of its own quickly through the neurocomputational mutations that can occur in cyberspace. It would be a mistake for governments to underestimate their speed or potency, or to overestimate their ability to predict or control these events.

How might such psychological tendencies be addressed to mitigate potentially harmful effects? Recall that, like with psychotherapy, awareness of a process does not necessary incur protection against its effect. Knowing that these biases exist does not prove sufficient to eliminate their effects. However, placing options side by side in ways which render transparent the effect of these biases at the time of decision can lead to more reliable choices. In addition, competence matters. Training leaders, especially those from cyber command who will be tasked with providing information to a decision maker in a time of crisis, with as much information as possible about both the technical and political aspects of attribution, may prove enormously helpful. In this regard, simulations where such leaders learn to more properly calibrate their judgments, and their confidence, with immediate feedback as to accuracy, can go a far distance toward increasing skill, decreasing overconfidence, and rendering more informed choice. Such procedures could be instigated as part of clear institutional promotion and training requirements for relevant jobs.

Prospects for cyberattack will not diminish. Developing appropriate response strategies in the face of various potential scenarios will prove challenging. But marshalling the forces of human psychology in support of such procedures constitutes a stronger edifice upon which to build defense than theories modeled on assumptions of rationality without empirical support.

REFERENCES

Bar-Joseph, Uri & McDermott, Rose. 2007. Personal Functioning Under Stress Accountability and Social Support of Israeli Leaders in the Yom Kippur War. *Journal of Conflict Resolution* 52(1): 144-170.

Bargh John & McKenna KYA. 2004. The Internet and social life. *Annual Review of Psychology* 55:573-90.

Bar-Hillel, Maya & Neter, Efrat. 1993. How alike is it versus how likely is it: A disjunction fallacy in probability judgments. *Journal of Personality and Social Psychology* 65(6): 1119-1131.

Carpenter, Daniel. 2002. Groups, the Media, Agency Waiting Costs, and FDA Drug Approval. *American Journal of Political Science* 46 (3): 490-505.

Cialdini, Robert. 2001. *Influence: science and practice, 4th ed*. New York: HarperCollins.

Derks, Daantje, Agneta H. Fischer, Bos, and E. R. Arjan. 2008. The role of emotion in computer-mediated communication: A review *Computers in Human Behavior* 24(3):766-785.

Ehrlinger, Joyce, Johnson, Kerry, Banner, Matthew, Dunning, David & Kruger, Justin. 2008. Why the Unskilled Are Unaware: Further Explorations of (Absent) Self-Insight Among the Incompetent. *Organ Behav Hum Decis Process* 105(1): 98-121.

Freud, Sigmund. 1922/1990. *Group Psychology and the Analysis of the Ego*. New York: W. W. Norton & Co.

Gault, Barbara & Sabini, John. 2000. The roles of empathy, anger, and gender in predicting attitudes toward punitive, reparative, and preventative public policies *Cognition & Emotion*, 14, (4):495-520.

Gigerenzer, Gerd. 1996. On narrow norms and vague heuristics: A reply to Kahneman and Tversky. *Psychological Review* 103(3): 592-596.

Guadagno, Rosanna, Cialdini, Robert, & Evron, Gadi. 2010. Storming the Servers: A Social Psychological Analysis of the First Internet War. *Cybertechnology, Behavior and Social Networking* 13: 1-8.

Heath, Chip & Heath, Dan. 2007. *Made to Stick: Why Some Ideas Survive and Others Die*. New York: Random House.

Jervis, Robert. 1979. Deterrence Theory Revisited. *World Politics* 31 (2): 289-324.
Jervis, Robert. 1982-1983. Deterrence and Perception. *International Security* 7 (3): 3-30.
Jervis, Robert. 1998. *Systems Effects: Complexity in Political and Social Life.* Princeton, NJ: Princeton University Press.
Johnson, Dominic. 2004. *Overconfidence and War*: The Havoc and Glory of Positive Illusions. Cambridge, MA: Harvard University Press.
Johnson, Dominic, McDermott, Rose, Barrett, Emily, Cowden, Jonathan, Wrangham, Richard, McIntyre, Matthew & Rosen, Stephen Peter. 2006. Overconfidence in wargames: experimental evidence on expectations, aggression, gender and testosterone. *Proc Biol Sci*. 273(1600): 2513-2520.
Johnson, Dominic & Fowler, James. 2009. The Evolution of Overconfidence. Unpublished ms.
Joinson AN. 2001. Self-disclosure in computer-mediated communication: the role of self-awareness and visual anonymity. *European Journal of Social Psychology* 31:177–92.
Kahneman, Daniel & Tversky, Amos. 1979. Prospect Theory: An Analysis of Decision under Risk. *Econometrica* 47 (2): 263-291.
Kahneman, Daniel & Tversky, Amos. 1984. Choices, Values and Frames. *American Psychologist*. 39(4): 341-350.
Kelley, Harold. 1967. Attribution Theory in Social Psychology. *Nebraska Symposium on Motivation* 15: 192-238.
Khong, Yuen Foong. 1992. *Analogies at War: Korea, Munich, Dien Bien Phu and the Vietnam Decisions of 1965.* Princeton, NJ: Princeton University Press.
Lerner, Jennifer & Keltner, Dacher. 2000. Beyond valence: Toward a model of emotion-specific influences on judgement and choice. *Cognition & Emotion* 14 (4): 473-493.
Lord, Charles, Ross, Lee & Lepper, Mark. 1979. Biased assimilation and attitude polarization: The effect of prior theories on subsequently considered evidence. *Journal of Personality and Social Psychology* 37 (11): 2098-2109.
MacKay C. 1841. *Extraordinary Popular Delusions and the Madness of Crowds*. New York: Farrar, Straus, & Giroux.
Matheson K & Zanna Mark. 1989. Persuasion as a function of selfawareness in computer-mediated communication. *Social Behaviour* 4:99–111.
Moral-Toranzo, Felix, Canto-Ortiz, Jesus, & Gomez-Jacinto, Luis. 2007. Anonymity effects in computer-mediated communication in the case of minority influence. *Computers in Human Behavior* 23 (3):1660-1674.
National Research Council. 2009. *Technology, Policy, Law, and Ethics Regarding U.S. Acquisition and Use of Cyberattack Capabiltities.* Washington, D.C.: National Academies Press.
Postmes T, Spears R, & Lea M. 2002. Intergroup differentiation in computer mediated communication: effects of depersonalization. *Group Dynamics* 6:3-15.
Reicher Stephen, Spears R, & Postmes T. 1995. A social identity model of deindividuation phenomena. In Stroebe W, Hewstone M, eds. *European review of social psychology, vol. 6*. Chichester: Wiley.
Rosen, Stephen Peter. 2005. *War and Human Nature.* Princeton, NJ: Princeton University Press.
Sassenberg K, Boos M, Rabung S. 2005. Attitude change in face-to face and computer-mediated communication: private selfawareness as mediator and moderator. *European Journal of Social Psychology* 35:361-74.
Slovic, Paul.1999. Trust, Emotion, Sex, Politics, and Science: Surveying the Risk-Assessment Battlefield. *Risk Analysis* 19(4):689-701.
Slovic, Paul, Kahneman, Daniel & Tversky, Amos. 1982. *Judgment Under Uncertaitny: Heuristics and Biases.* New York: Cambridge University Press.
Stoll, Clifford. 2005. *The Cuckoo's Egg; Tracking a Spy Through a Network of Computer Espionage.* New York: Pocket.
Tversky, Amos & Kahneman, Daniel. 1983. Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review* 90(4): 293-315.
Tversky, Amos & Kahneman, Daniel. 1992. Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty* 5(4): 297-323.
Webster, Donna & Kruglanski, Arie. 1997. Cognitive and Social Consequences of the Need for Cognitive Closure. *European Review of Social Psychology* 8.
Zimbardo, Philip. 1970. The human choice: individuation, reasons, and order versus deindividuation, impulse, and chaos. In Arnold WJ, Levine D, eds. *Nebraska Symposium on Motivation* 17:237-307. Lincoln: University of Nebraska Press.