Toward Practical Cyber Counter Deception

Christopher Porter

Nation-states increasingly engage in strategic deception in cyberspace, frustrating traditional counter deception approaches. This paper evaluates and critiques the philosophical underpinnings and practical implications of existing military-political counter deception and computer forensic approaches. Analysts can better detect and expose strategic deception campaigns in cyberspace by focusing on the size and organizational strength threat actors need to conduct the operations.

Christopher Porter is the manager of Horizons, the strategic intelligence and forecasting arm of the cybersecurity company FireEye. He previously served for nine years at the CIA. This article does not necessarily represent the views of the U.S. government.

Journal of International Affairs, Winter 2016 Vol.70, No.1.

Detecting, countering, and deterring strategic deception in cyberspace remains reliant upon techniques and policies developed for countering deception in the physical world. Solid assumptions in a resource-constrained physical space are largely inapplicable to forensic examination of cyberspace, where resources are effectively limitless. Specifically, counter deception methods used by military and intelligence officers rely on the assumption that would-be deceivers either leave behind evidence incongruous with the reality they are attempting to present or incompletely simulate the physical properties of the reality they are attempting to mimic. Cyber threat actors, some probably sponsored by the Russian government, often exploit the reliance on these assumptions, as physical counter deception techniques do not apply to cyberspace. Ironically, the more trained and experienced an analyst is in detecting deception, the more ill-suited they may be to detecting cyber deception using current methods and training. Longstanding difficulties attributing cyber operations to a particular nation-state sponsor, compounded by a lack of reliable counter deception tools, have elevated cyber deception to a politically effective weapon unto itself.

This paper highlights the challenges inherent in cyber counter deception and provides specific questions that analysts can ask to overcome these challenges, particularly as part of an analysis-of-competing-hypotheses structured analytic exercise. The paper concludes with a short discussion of the importance public education could play in mitigating the effectiveness of cyber deception, as currently practiced when targeting democracies.

Meatspace Counter Deception Methods Do Not Hold Up in Cyberspace

Barton Whaley has often been called the father of modern deception studies. So extensive is his influence over deception and counter deception scholarship that the Deception Research Center at CIA headquarters bears his name. Perhaps the most unexpected data source Whaley drew upon was the work of stage magicians. He studied the cognitive biases being exploited by magicians and their frequency across many tricks. The "second rule of every deception's weakness," credited jointly to Whaley and Jones, is as follows: "Creating a deception simultaneously creates all the clues (incongruities) needed for its detection." Whaley's own corollary to that rule is that "every deception creates at least two incongru-

ities—one about the thing being hidden (dissimulated), the other about the thing being shown (simulated) in its place."[1] Although strategic counter deception practitioners do not directly consult his studies as a standard of care, Whaley presents ample evidence that the general trends he identifies are reasonable proxies for the frequency that various military and strategic political deception methods are employed.

Table 1: The Bell-Whaley Matrix of Deception (as applied to 60 magic tricks)

	Simulating		
Dissimulating	Mimicking	Inventing	Decoying
Masking	19	10	10
Repackaging	4	3	1
Dazzling	10	2	1

The matrix of deception is arrayed in order of likely success, such that a trick that masks and mimics is thought to be more likely to succeed and remain in use than one that relies on dazzling and decoying.

- ***Masking*** hides the real by making it invisible.
- ***Repackaging*** hides the real by disguising.
- ***Dazzling*** hides the real by confusing.
- ***Mimicking*** shows the false by having one thing imitate another.
- ***Inventing*** shows the false by displaying another reality.
- ***Decoying*** shows the false by diverting attention.

These deceptions, and the means by which an audience member or fellow conjurer could detect them, were then applied to great military and strategic planning problems across history. While this line of inquiry bore a lifetime of fruit for Dr. Whaley, it rests on one uncomfortable premise: that the human mind's errors in processing physical phenomena like sight and sound can lead to strategic miscalculation, thereby opening the door to deception. Whaley found that, because of human weakness in discerning the direction from which sound is coming, many magicians' tricks relied on deceiving human hearing. However, there is no clear analogue to cyber operations for this observation, therefore,

the data he collected may not give a representative sample of cyber deception operations. His work speaks for itself in terms of accurately describing historical military deception and useful application to myriad intelligence problems of the present (particularly for clandestine operatives), but the parallels it draws fall short in today's most important theater of deception: cyberspace.[2] This suggests that mistakes are being made writ large by national security practitioners attempting to apply concepts learned from historical military deception studies to a non-physical space.

Barton Whaley did consider the application of computers to counter deception. Building on the work of R.V. Jones's studies in practical joking and information processing, Dr. Whaley theorized that computers, because they are good at detecting incongruities, could be applied as "expert systems" to alert counter deception practitioners of suspicious patterns. Yet, despite the wide availability of increasingly advanced intrusion detection systems that automatically discover anomalous behavior on computer networks, cyber deception at the strategic level remains rife and is increasingly effective in many ways.

Even a cursory examination of cyber operations reveals that the proportions found in magicians' tricks and, according to Whaley, on the battlefield are warped compared to what one finds in the practice of counter deception in cyberspace.[3] For example, any stateful connection (Web traffic, e-mail, etc.) will require a two-way connection of some sort in order to succeed. While cyber operations can endeavor to hide the purpose or maliciousness of their traffic, the traffic itself is not ordinarily masked or hidden in the sense that Whaley means here. On the other hand, repackaging of malicious software and operations has increased and become so common as a means of compressing malicious software inside seemingly legitimate code that the term "packer" became a term of art in comparison to the relative infrequency of repackaging in the deception matrix.[4]

Similarly, the relatively rare use of inventing and decoying does not match up with practical cyber operations experience, where the creation of false fronts to carry out operations is frequently plausible and the common use of "honeypots" to entrap would-be attackers in dummy networks suggests that the low percentages observed in magic tricks do not apply in cyberspace.[5] Mimicking,

meanwhile, is still useful but may be less necessary than in physical operations, given the unreliability of forensic cyber indicators as a means of attribution, as described below.

The frequency we would expect to find these methods of deception in cyber operations is out of proportion, because the tricks of magicians and field generals rely on physical constraints that do not apply to cyberspace. While it may be resource intensive to create false, blow-up aircrafts as decoys, their cyber equivalent requires no resources to be consumed beyond keystrokes. Repackaging and inventing also face resource constraints in the physical world that simply do not exist in the nonphysical realm where obfuscated computer operations remain costly but feasible at the scale needed for physical deception operations. Dazzling, because it operates so directly on the mind by generating noise to drown out useful signal, is particularly applicable to the cyber realm where the material cost of sending signals is already low in much of the developed world and rapidly falling globally.[6]

The practical effect of this change in deception likelihood, particularly for the traditionally rare combinations, is that the intelligence analyst investigating a hypothesis of deception in the case of a major cyber operation is likely to do so from a position with a bad estimate of that deception's likelihood.

The Mind, Not the Senses, Are the Hackers' Ultimate Deception Target

Almost all of counter deception is premised on something similar to Whaley's rules being true, leading to the conclusion that deception by an adversary can always be detected because an environment in which strategic deception is taking place will always have additional forensic artifacts that do not belong and additional pieces of evidence that one would expect to find but are absent.

When applied to cyberspace, the first premise would imply that cyber deception operations must necessarily leave behind forensic evidence that would point to the deception's existence, particularly if investigators were clever, quick, and cooperative enough to uncover them. However, this would be a misunderstanding of such investigations, which result not in raw physical evidence but in analysis

of evidence, more akin to a blood spatter report but not the blood itself.

These basic elements of cyber deception exploit the philosophical underpinnings for Whaley's rules—the philosophical underpinning for which he draws from the work of criminologist Edmond Locard, whom he credits for making "the single greatest contribution…to the theory of detection" in what is known as Locard's exchange principle. The principle states that the "perpetrator leaves at the scene marks of his passage; and, by an inverse action, simultaneously takes away on his body or clothes indications of his visit."[7] Dr. Whaley expounds on the theory, stating that the principle "applies to all types of deception, although we must add *psychological* perceptions to Locard's *physical* evidence." Cyber deception operations, as we saw above, need not involve physical evidence at all, and the detection apparatuses being used are without psychologies to speak of, casting into doubt the premise that cyber deception operations would necessarily leave behind incongruous evidence. But an additional, fundamental criticism is that cyber investigative work does not discover raw evidence as the traditional criminologist means it; it only discovers interpretations, which can neither be taken away nor left at the scene because they exist in the investigator's mind.

Following a cyber attack or exfiltration of sensitive data as the result of a cyber operation, attempts to attribute the incident to a particular actor inevitably build upon the forensic evidence used by specialists called in to remediate the intrusion and remove unwanted attackers and malicious software from the targeted network. The reports are inherently interpretive, as they describe logical actions being carried out on a computer system or network, rather than the physical means by which they happen. If a file is moved, is that a legitimate or a dangerous action? What if a file is deleted? What if new information is created? The actions of the software do not, unto themselves, have either a legitimate or malicious tint. Only in context, with interpretation by a potentially biased human being, do the actions connote a threat.

Behavior-based assessments, while highly reliable and very useful for detecting and removing intruders, are scientific but do not deal with immutable physical properties. For example, one of the basic actions of cyber intruders is to delete or modify the very logs recording their actions. Also, the infrastructure used to launch cyber intrusions is quite often a machine belonging to a previous victim,

creating a recursive hall of mirrors that permits attackers to not only avoid attribution to their sponsors but to create misleading forensic evidence for investigators to find.[8]

Would-be cyber deceivers can use malicious software that remains resident in computer memory, rather than written to disk, leaving behind little to no evidence to find. Still, other intruders use strong encryption to scramble the evidence on their victims' machines, leaving little to be detected after the fact unless separate, specialized equipment records network and computer activity to specifically defeat these techniques. Even then, most advanced cyber intruders now use legitimate Web sites, like social media, discussion boards, and news Web sites to launch their attacks.[9] Rather than creating advanced cyber tools, exploiting built-in system components leaves little trace that concretely points to a particular actor.[10] The investigative methods of detection still make defense and expulsion possible, but they are not definitive indicators of the guilt of any one party in the same way that DNA evidence or fingerprints might be.

Importantly, cyber deceivers appear to know the above truths already and appear to have internalized them better than most network defenders and almost everyone involved in cyber counter deception. As a whole, cyber threats have migrated from trying to avoid leaving evidence that absolutely incriminates them, to trying to avoid detection, to the current situation where it is common for sufficiently advanced actors to leave behind deliberately misleading evidence that points investigators' "forensic analysis" toward another party's culpability in "false-flag" attacks.[11]

One of the most high-profile of these events was the attack on Sony Pictures Entertainment in 2014, which the FBI publicly attributed to North Korean sponsorship.[12] Assuming the FBI is correct, Pyongyang did not try to avoid blame for the attack. Their cyber warriors did not appear as "ghosts in the system." Rather, they gave credit to an apparently fictitious non-state hacking group, the Guardians of Peace. Even though some technical indicators pointed to North Korean networks as the origin, and political events at the time made them the only plausible state sponsor, many public experts and even some cybersecurity companies were skeptical that the attack was carried out by North Korea. The group was immature, most of its demands were related to causes only Western audiences

would care about, and any forensic evidence pointing to North Korea could easily be explained as the work of actors trying to frame North Korea. Forensic evidence is so widely held to be insufficient for attribution that attackers need not even necessarily hide such evidence. The mere suggestion it could be easily hidden or spoofed is enough to cast attribution into doubt. In democracies, such doubt has repeatedly led to policymaking paralysis, as national leaders struggle to convince their publics that responses to damaging cyber operations strong enough to deter future such actions are appropriate.

Bureaucracy is the Dog That Barks

If we cannot trust our "lying eyes" or at least a reasonable interpretation of what we think we see, then what can we trust? Remember that in the principles of counter deception, those attempting to discern a deception—in this case, to detect the true sponsor of a cyber operation rather than the cover—have two bites at the apple: to discover incongruous evidence left at the scene, which we have shown is not reliable and may even be a folly exploited as part of the deception, and to discover incongruities in the effect being shown in its place. It is here that I believe hope remains for detecting the best cyber deceivers: finding the lack of evidence for an ostensible actor's involvement and incongruities in organizational complexity that arise from the hidden sponsor's use of cover. As Sherlock Holmes found, it is sometimes the dog that does not bark in the night that is the most peculiar thing.

Dr. Whaley quotes approvingly of the work of Dr. J. Bowyer Bell and his observations on cognitive asymmetry as applied to modern terrorist groups.[13] Of particular relevance to this discussion is Dr. Bell's finding that "all undergrounds are inefficient…those that survive at all—appear awesome, effective, but each must operate within an ecosystem that may permit persistence but at great cost in efficiency." He goes on further to say that the "more secret an organization, the more inefficient," and that all "rebels are, to a considerable degree, impractical and incompetent."

If we take the stereotype of non-state hackers as rebels to its natural conclusion, we ought then to wonder about the most damaging cyber data theft operations and attacks. Where did such disparate cyber rebels get their bespoke tools? More

usefully, how did they organize themselves and maintain persistence on the target network and a full work schedule in what turned out to be an attack that yielded no financial profit? Who funded their operation? If the attackers are ostensibly a non-state group, where is the evidence on forums and the dark Web of their prior activities and hacking targets?

Simply put, the larger an organization would need to be to sustain a cyber operation over time, with multiple methods and specialties and sustained effort over time, the more likely that a large, well-organized group rather than a "rebel" from the "underground" is responsible. The precise technical means, forensic indicators, and claims of responsibility are all ancillary to this, much-harder-to-disguise sign.

For example, the Russia-backed Advanced Persistent Threat (APT) group known as APT29 gained initial access to many of its Western government victims by sending carefully crafted emails, which in a different context could easily be mistaken for the common spam of petty cybercriminals worldwide.[14] The group mostly eschews the use of their armory of zero-day exploits (previously undiscovered bugs that advanced attackers can exploit with sometimes-significant resource investment), in favor of using built-in system components and hijacking legitimate Web services and social media profiles—all techniques available to individual hackers.

Yet, a review of the pace of their operations, speed at recovering from incident remediation, and careful targeting of key technical personnel make it clear that the level of organization behind their operation is far greater than any individual, hacktivists collective, or less sophisticated nation-state actor. Examining other evidence through this lens, like APT29's focus on European governments, Western national leaders, and topics related to the Russia–Ukraine conflict, along with forensic evidence, increase analytic confidence in saying that the group operates at Moscow's behest.

Without understanding their operational pace and lateral targeting prowess, and hence their organizational size and sophistication, it would have been easy to dismiss these incidents as the work of one of an endless string of Russia-based cybercriminal groups. Moscow has for the past year been increasingly aggressive

in its cyber operations, frequently persevering against corporate and political targets even once its operations are detected and eschewing the use of more deniable infrastructure and novel tools.

Indeed, some of its most prominent operations such as the exploitation of the Democratic National Committee and Democratic Party donors and APT28's disruption of French television station TV5Monde, acting under the guise of a pro-Islamic State hacktivist group, seem to have been designed to avoid detection only long enough for operations to successfully conclude.[15] Those and other operations increasingly leave behind forensic evidence with clear links to the Russian government, despite those groups' demonstrated ability to control operations in highly deniable ways, such as by using commercial satellites, in what is likely a deliberate calculation that a patina of deniability is sufficient deception.[16] Likewise, the agility with which information compromised via APT28's espionage operations is turned into effective propaganda and spread on Russia-linked social media and news outlets implies sponsorship by an organization much larger than the few on-keyboard operators necessary to control only the exploitation of individual targets.

In more recent cases, such as the compromise of current and former senior U.S. leaders and politicians and subsequent release of embarrassing personal e-mails to the public, APT28 chose to forego many of its better methods for complicating attribution and probably left behind some forensic artifacts, such that experts, but not the general public, would quickly attribute their activities to Russian state sponsorship as a form of bilateral messaging by the Kremlin. U.S. policymakers and forensics experts who see Moscow's hand in these acts are increasingly left at odds with the general public, whose attitudes, guided by press coverage that gives undue credence to minority reports, are ambivalent to the point of causing policy paralysis.

Giving national security decisionmakers and, often equally as important, the public a better sense of the probability that a given major cyber incident was conducted by a nation-state soon after it is discovered could be key to improving cyber response in democratic nations. As much as we may want to believe the underdog narrative, it is most often the case that the better organized and prepared player will win a security engagement. In any event, even an attacker

that is successful in gaining access to a target will need to be organized to exploit their access for any real gain. In serious cyber incidents, a focus on the organizational level of effort will bring to light the incongruities needed to uncover cyber deception. While not a catch-all, a good first set of questions to ask when deception is possible as part of a major cyber operation should include:

- What level of organization would be necessary to target, exploit, use, and fund the activity behind this operation? Examining further forensic evidence and geopolitical intelligence through this lens would help to detect the presence, though not necessarily the culprit, behind cyber deception.
- How many people would need to be involved, while maintaining secrecy, in order to carry out this operation at the pace it is happening? The more people and the greater secrecy needed, the less likely it is that disparate actors are self-organizing across the globe and the more likely it is that some central bureaucracy, like a state intelligence or military apparatus, is managing their actions in concert.
- Which organization has better access to cyber technology? If it is the victim, the better their cyber savvy and investment in defense, the less likely a group "punching up" is likely to have really done so, at least on their own.

Naturally, some cyber threat groups will be able to compromise targets of opportunity, use more advanced tools, or have more experience than others in ways that require different levels of organization and staffing. Applying these questions as part of a formal "analysis of competing hypotheses" that seeks to eliminate improbable cover operations may therefore be useful at highlighting incongruities even when the answers do not provide definitive positive evidence on their own.[17]

Better Cyber Policy Could Aid Counter Deception

Dr. Whaley notes that deception fails under "only four circumstances, when the target:

- takes no notice of the intended effect,

- notices but judges it irrelevant,
- misconstrues its intended meaning, or
- detects its method."[18]

The first circumstance is, by definition, unlikely in the event of major cyber attacks, since the effect will almost certainly be noticed. Detecting the method of deception is the realm of technical and intelligence experts, the tradecraft of which I hope this article has advanced. Though, the opportunity remains to reduce cyber deception not by intelligence methods but by altering the playing field such that would-be deceivers have a harder time delivering their message, such as misdirecting blame for a cyber attack they conduct.

One of the best ways of reducing the plausibility of cyber deception would be to better educate the press and the public at large that large organizations, such as advanced states and organized crime groups, use smaller organizations, like hacktivist groups and individual criminals, as cover for their cyber operations. As we discussed above with respect to the Sony case, the success of cyber deception and resulting complications for attribution have become so widespread that many attackers do not even feel the need to engage in covering some of their tracks, since plausible deniability will exist either way. Raising public skepticism regarding false-flag attacks would go a long way toward reducing their attractiveness in the first place.

It is also worth examining what techniques will not work. Rapid disclosure of new vulnerabilities, threat actor infrastructure and techniques, indicators for cyber espionage, and attack tools are held up as the gold standard for network defenders. While this approach has been useful in protecting organizations from financially motivated non-state cybercriminals, the most advanced adversaries seem to be significantly less affected by these activities. A study by my colleague Kristen Dennesen noted that APT28—often rumored in press reporting to be sponsored by Russia's Main Intelligence Directorate (GRU)—was the subject of more than 20 industry reports from October 2014 to October 2015 examining the group's tools and tactics, yet they "sustained operations in spite of near-constant public reporting."[19] Chinese groups like APT3 and APT17 have likewise shown the ability to retool operations within hours of exposure, casting doubt on the efficacy of information sharing-based defenses to disrupt well-resourced

cyber operations conducted with the protection of strategic deception.

However, reducing the volume of sophisticated cybercrime would make such attacks less politically viable. Currently, many governments and network defenders simply cannot "see the forest for the trees" well enough to marshal public opinion against the suspected offenders. Burning down some of the lesser trees by treating cybercrime more seriously as a national security, rather than a law enforcement, responsibility would naturally leave less noise for cyber deceivers to hide their signal.

Endnotes

1 Barton Whaley, *Practise to Deceive: Learning Curves of Military Deception Planners* (Annapolis, MD: Naval Institute Press, 2016).
2 Miles A. McQueen and Wayne F. Boyer, "Deception Used for Cyber Defense of Control Systems" (report, Idaho National Laboratory, 2009).
3 Quantitative analysis of the frequency of deception in cyber operations in comparison to those of the Bell-Whaley Matrix of Deception remains a promising avenue of future research. Alas, collection bias for such research also remains as a formidable barrier.
4 Fanglu Guo, Peter Ferrie, and Tzi-cker Chiueh, "A Study of the Packer Problem and Its Solutions" (report, Symantec Research Laboratories, 2008).
5 Thomas Rid and Ben Buchanan, "Attributing Cyber Attacks" *Journal of Strategic Studies* 38, no. 1-2 (2015).
6 Scott Kipp, "Exponential Bandwidth Growth and Cost Declines," *Network World*, 10 April 2012.
7 Edmond Locard, *L'Enquête Criminelle et les Méthods Scientifiques* (Princeton University, 1920).
8 The ease with which actors could implicate mutual adversaries for cyber attacks against one another and lead them into conflict highlights what a terrible idea "hacking back" and other mutually assured destruction-like schemes would be with present technology. Such policies are a cyber deceiver's dream.
9 "Hiding in Plain Sight: FireEye and Microsoft Expose Obfuscation Tactic" (report, FireEye, May 2015).
10 "HAMMERTOSS: Stealthy Tactics Define a Russian Cyber Threat Group" (special report, FireEye, July 2015).
11 Michael Mimoso, "APT Attackers Flying More False Flags Than Ever," *Threat Post*, 17 March 2016.
12 FBI Press Release, "Update on Sony Investigation," 19 December 2014.
13 J. Bowyer Bell, "Dragonworld (II): Deception, Tradecraft, and the Provisonal IRA," *International Journal of Intelligence and Counterintelligence* 8, no. 1 (1995), 21-50.
14 "HAMMERTOSS: Stealthy Tactics Define a Russian Cyber Threat Group," 9.
15 Dmitri Alperovitch, "Bears in the Midst: Intrusion into the Democratic National Committee" (blog, CrowdStrike, 15 June 2016); "TV5 Monde attack 'by Russia-based hackers,'" BBC, 9 June 2015.
16 Max Fisher, "Why Security Experts Think Russia Was Behind the D.N.C. Breach," *New York Times*, 26 July 2016; Patrick Tucker, "Russia Wanted to Be Caught, Says Company Waging War on DNC Hackers," *Defense One,* 29 July 2016.
17 Richards J. Heuer, Jr., *Psychology of Intelligence Analysis* (Center for the Study of Intelligence, CIA, 1999).
18 Barton Whaley, *Textbook of Political-Military Counterdeception: Basic Principles & Methods* (National Defense Intelligence College, 2007), 99.
19 Kristen Dennesen, "Hide and Seek: How Threat Actors Respond in the Face of Public Exposure" (remarks, SANS Cyber Threat Intelligence Summit, San Francisco, CA: 4 February 2016).