NPS-DA-14-001

NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

CYBER ANALOGIES

Edited by

Emily O. Goldman and John Arquilla

February 28, 2014

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From-To)*
28-02-2014	Technical Report	2013-2014

4. TITLE AND SUBTITLE

Cyber Analogies

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Editors: Emily Goldman and John Arquilla

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) AND ADDRESS(ES)
Department of Defense Analysis
Naval Postgraduate School,
Monterey, CA 93943

8. PERFORMING ORGANIZATION REPORT NUMBER
NPS-DA-14-001

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

US Cyber Command
9800 Savage Road
Fort Meade, MD 20755

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public release; distribution is unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

This anthology of cyber analogies will resonate with readers whose duties call for them to set strategies to protect the virtual domain and determine the policies that govern it. Our belief it that learning is most effective when concepts under consideration can be aligned with already-existing understanding or knowledge. Cyber issues are inherently tough to explain in layman's terms. The future is always open and undetermined, and the numbers of actors and the complexity of their relations are too great to give definitive guidance about future developments. In this report, historical analogies, carefully developed and properly applied, help indicate a direction for action by reducing complexity and making the future at least cognately manageable.

15. SUBJECT TERMS
Cyber analogies, Cyber, US Cybercom, Cyber Pearl Harbor, Cyber warfare, Cyber security

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			Dr. John Arquilla
Unclassified	Unclassified	Unclassified	UU	132	**19b. TELEPHONE NUMBER** *(include area code)* 831-656-2097

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18

THIS PAGE INTENTIONALLY LEFT BLANK

NAVAL POSTGRADUATE SCHOOL
Monterey, California 93943-5000

Ronald A. Route
President

Douglas A. Hensler
Provost

The report entitled "*Cyber Analogies*" was prepared for U.S. Cyber Command, 9800 Savage Road, Fort Meade, MD 20755 and funded by the department of Defense Analysis at the Naval Postgraduate School.

Further distribution of all or part of this report is authorized.

This report was prepared by:

Emily Goldman
Senior Advisor, U.S Cyber Command

John Arquilla
Chair, Defense Analysis

Reviewed by:

John Arquilla
Chairman
Department of Defense Analysis

Released by:

Jeffrey D. Paduan
Dean of Research

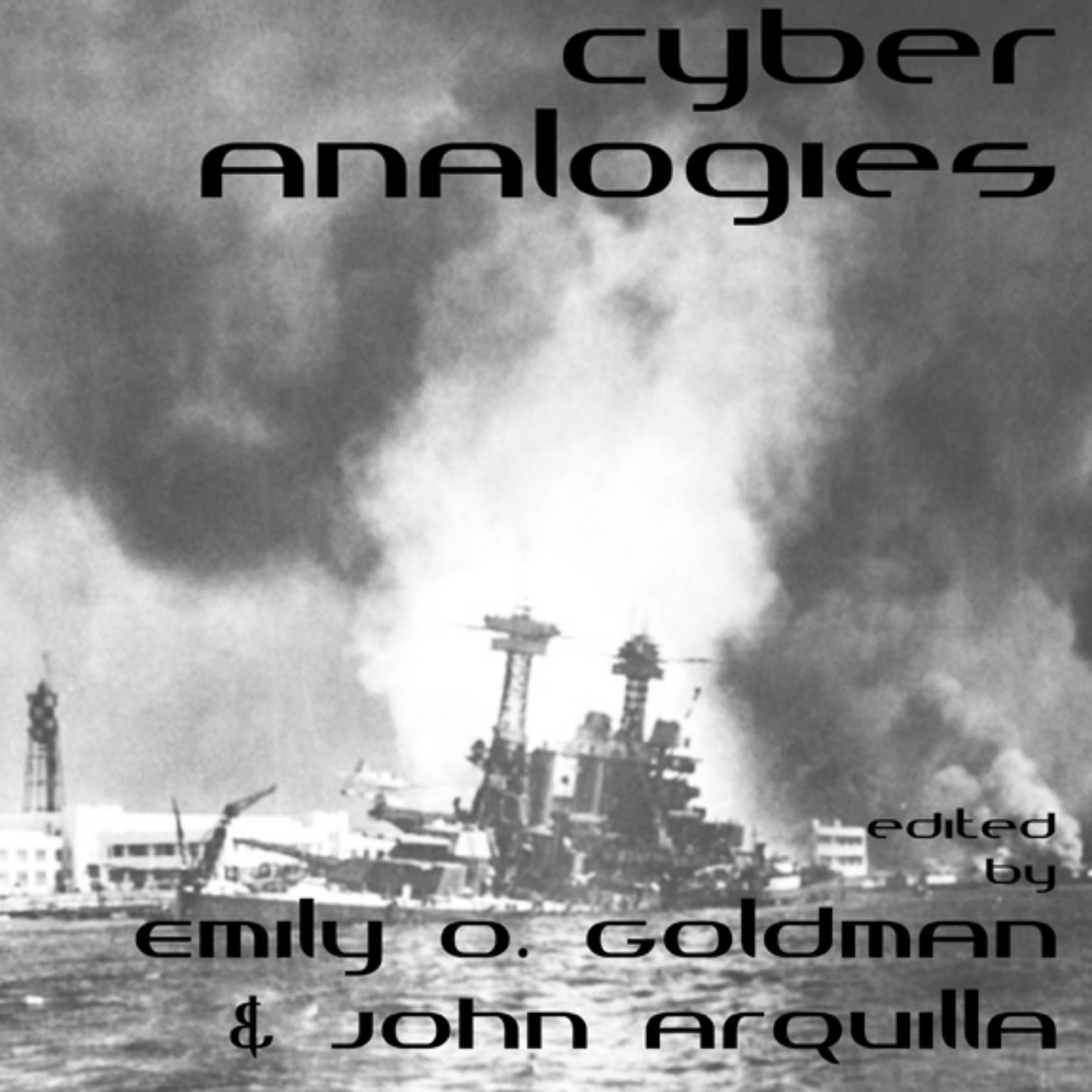
THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

This anthology of cyber analogies will resonate with readers whose duties call for them to set strategies to protect the virtual domain and determine the policies that govern it. Our belief is that learning is most effective when the concepts under consideration can be aligned with already-existing understanding or knowledge. Cyber issues are inherently tough to explain in layman's terms. The future is always open and undetermined, and the number of actors and the complexity of their relations are too great to give definitive guidance about future developments. In this report, historical analogies, carefully developed and properly applied, help indicate a direction for action by reducing complexity and making the future at least cognitively manageable.

THIS PAGE INTENTIONALLY LEFT BLANK

cyber analogies



edited by

emily o. goldman & john arquilla

This Technical Report was sponsored by United States Cyber Command, and produced under the aegis of the Department of Defense Information Operations Center for Research at the Naval Postgraduate School. The views expressed herein are those of the contributors. For reprint permissions or copies, please contact:

Rebecca Lorentz, Administrator
DoD IO Center for Research
Naval Postgraduate School
589 Dyer Road, Room 102
Monterey, CA 93943
rdlorent@nps.edu
831.656.7788

Technical Report: NPS-DA-14-001

TABLE OF CONTENTS

Introduction

The Cyber Analogies Project was launched in 2012 to assist U.S. Cyber Command in identifying and developing relevant historical, economic, and other useful metaphors that could be used to enrich the discourse about cyber strategy, doctrine, and policy. The intent of the project is to provide useful insights, both for those with little technical background in or direct connection to cyberwar and cyber security and for those whose job it is to think about the spectrum of cyber-related issues every day. The project was conceived and carried out to help very senior, busy, responsible people understand topics and issues that are fast-moving and dynamic, and have potentially great consequences for society, security, and world affairs.

The President has identified the cyber security threat as one of the most serious we face as a nation. Events in cyberspace continue to accelerate, as both nation-states and non-state actors seek to exploit asymmetrical advantages in this virtual domain. The risks and costs of failing to act skillfully and effectively are likely to be widespread, cascading, and almost surely highly disruptive. Many small-scale attacks can be—and are being—prevented; but a major, mass-disruptive attack is a growing possibility against which current defenses are inadequate.

There are several reasons for saying this. First, the nation's cyber architecture relies on a multitude of asset types, configurations of those assets, and multiple organizations pursuing strategies and policies—not always in concert. Second, there is no shared situational awareness across all networks, the absence of which minimizes chances of detecting adversarial activities. Third, the authorities, policies, rules of engagement, and division of effort to act in defense of cyberspace are not fully articulated, implemented, or even properly delegated. Fourth, there are insufficient trained and ready forces to act. Fifth, the information systems architecture has been built for availability, functionality, and ease of use, with considerations of security all too often an afterthought. Sixth, commercial firms in the private sector are reluctant to divulge penetrations, as this would acknowledge vulnerability and possibly cause a loss of customer and shareholder confidence. Finally, and most closely related to the military sphere, cyberwar has not yet been fully integrated into more traditional concepts of operations. These gaps increase vulnerability to malicious activity—perhaps even to mass disruption of critical systems in the physical world—coming from cyberspace.

The cyber threat continues to evolve swiftly. Things we value—personal wealth, national economic prosperity, intellectual property, national security secrets—are all targets. More and more, these treasures reside in or depend upon cyberspace—the new battleground or, in Pentagon parlance, the "warfighting domain where adversaries are already operating." Ability to keep pace with the cyber evolutionary curve, or perhaps to stay a step ahead of those who wish to do harm, depends on the ability to visualize how to fight in this new domain, just as strategists from earlier eras had to imagine how operating in the air, or with nuclear weapons, changed military affairs. Analogies drawn from earlier eras and different disciplines have the potential to help with visualization, allowing us to think through new or difficult subjects. They offer us an inductive approach to developing an understanding of high-level conceptual and operational issues related to cyber security and cyber warfare. Scrutiny of historical cases, which makes up much of this anthology, can at the very least lead to identification of similar problems and ways to address them—that either "worked" or did not. No historical example provides a perfect fit for current challenges, so any analogy must be carefully developed and not used in isolation. Still, historical analogizing is essential to the learning process. As H.G. Wells put it, "History is a race between education and catastrophe."

SURPRISE ATTACK

The primarily history-based analogies contained herein begin with a re-examination of Pearl Harbor. Teasing out implications of this case can help illuminate how to deal with the threat of surprise attack from cyberspace. Pearl Harbor remains

seminal, as the goal of the Japanese attack was not only to destroy key naval forces, but also to disrupt the ability of U.S. forces to respond to other strikes being launched simultaneously, across thousands of miles. The three chapters that begin this anthology (by James Wirtz, Michael Goodman, and Emily Goldman et al.) are cognizant of this operational-to-strategic linkage, and all derive key insights applicable to the process of shoring up cyber defenses against surprise attack. One of the most interesting observations is that Pearl Harbor was not a "bolt from the blue," but rather came at the darkest, lowest point in U.S./Japan relations. Warnings were issued, but the actions taken in response to these warnings made the Pacific Fleet more vulnerable to an air attack.

This analogy also suggests that we should not view cyber as separate from conventional military actions. A thinking adversary will probably want to link cyber with physical attacks. Further, opponents are likely to grow more interested in employing cyber surprise when they see it as a way to disable physical military capabilities. Historically, weaker parties turn to surprise to overwhelm stronger adversaries and prevent them from responding. Cyberwar techniques provide an opportunity to do exactly this, because modern militaries have grown so dependent on their information infrastructures. Indeed, the cyber domain holds the key to enabling—and disabling—logistics, as well as overall command and control.

The Pearl Harbor case—along with others examined in this anthology—suggests that the solution to developing an effective response in real-time is not entirely a technological one. There is also a strategic response, guided by answers to the question "What kind of national strategy would deter the opponent from launching an attack in the first place?" The best way may be to convince the adversary that a cyber Pearl Harbor will not achieve the desired effects; that the vulnerabilities they perceive as great are really much less so. The key is to reduce the attacker's confidence that conventional capabilities can be crippled by strikes on cyber targets. And if there is a reluctance to escalate, i.e., to retaliate for a cyber attack with a "kinetic" response, then mitigation of the effects of the former becomes a key to deterrence, convincing the adversary that he is not winning. Currently, if we have a deterrence/mitigation strategy, it is based on preserving uncertainty in the adversary's mind. If we can be more overt about mitigation, then we can add to the realm of uncertainty the factor of risk, so that the opponent is forced to ask the uncomfortable question: Which do we prefer: uncertainty or risk?

NUCLEAR PRE-DELEGATION

Given that avoiding another Pearl Harbor—i.e., a crippling "bolt from the blue" attack—was a prime driver of strategic thought in the nuclear era, the cyber analogies team thought it important to revisit this period, mining it for insights. Peter Feaver and Kenneth Geers have done just this, focusing on the issue of pre-delegated authority to act, given the swift tempo of operations likely to unfold. This was a very significant problem when it came to dealing with the massive destruction that would accompany nuclear attacks delivered by inter-continental missiles that could reach anywhere across the world in about half an hour. In cyberspace, where the effects are likely to be simply disruptive—at least for now—the attacking timeline, and thus the need to respond, can be reduced to seconds. Pre-delegated authority is clearly going to be an important issue area into which strategists and policy makers will have to wade deeply. The analogy from the nuclear era is no doubt going to add value to their discourse.

Pre-delegation raises the possibility that certain threat scenarios may oblige the national command authorities to do something that they would actually prefer not to—i.e., craft a set of allowed automatic military responses prior to any conflict. Pre-delegation was an inexpensive response to the Cold War nuclear "tri-lemma" (nukes must always be available for use; must never be used under conditions political leadership does not intend; and must remain under civilian control). Although a cheap solution, pre-delegation raised—and continues to raise—tricky questions about implementation: How far down the chain of command should pre-delegation go? How public can pre-delegation protocols be? Can they be reversed? Under what circumstances and to which systems does pre-delegation apply? What is the role of pre-delegation in terms of signaling to and establishing credibility with potential attackers? What does pre-delegation mean for private sector actors, given that, unlike nuclear weapons, there is no consensus that cyber is intrinsically a governmental function?

There are many skeptics—and many with much credibility—who believe that cyber attacks are not now serious enough to merit pre-delegation. This skepticism must be addressed before any decisions about pre-delegation can be made. Eventually, a situation may emerge wherein pre-delegation is seen as so obvious and necessary—because of clearly demonstrated advances in cyber disruptive capabilities—that operating without it becomes unimaginable. Thus the nuclear age provides an important example to guide thinking on this topic. Early on, military custody of nuclear weapons was hugely controversial. The initial idea was that civilians would physically hold the nukes to ensure civilian control. Pre-delegation was almost unthinkable. At some point, though, the decision was made that the military needed to retain physical custody. Operational issues surrounding pre-delegation did not go away, but the controversy did. What was considered unthinkable in 1946 had become obvious by 1966. But we are still "in 1946," politically and psychologically, on the issue of cyber pre-delegation.

AIRPOWER AND AIR DEFENSE

Perhaps the historical analogy that comes closest to explaining cyberwar is offered by the rise of airpower a century ago. From very early on, strategic thought about attack from the air diverged along two paths: one emphasized striking directly at an enemy homeland, without first having to defeat opposing land and naval forces; the other focused on the use of attack aircraft in close support of ground forces, or fleets at sea. Gregory Rattray's contribution to this anthology keeps the dual nature of airpower clearly in mind, and provides an opportunity to consider whether strategic thinking about cyberwar reflects an awareness that this duality applies to the virtual domain as well. The history of air campaigns over the past century suggests that strategic attacks have, for the most part, failed to achieve the material or psychological effects desired. But airpower used in close support of military and naval forces completely transformed the face of battle in the 20th century. If the air analogy holds, then the possibility is that strategic cyberwar—so greatly feared today—may have less impact on strategic affairs than cyber attacks mounted in conjunction with other military operations.

With this in mind, Rattray hypothesizes that, as was the case with airpower, the attacker might have an advantage; but the swift pace of technological change might reverse this in favor of the defender. His discussion of the defeat of the Luftwaffe in the Battle of Britain—so soon after its run of *blitzkrieg* victories—is particularly illuminating. Another key topic that Rattray considers closely is the matter of human capital. In the realm of airpower this has always been a crucially important factor; in the cyber domain the need for highly skilled, outstanding personnel is just as critical, if not more so. The implications for cyber education and training are profound, especially given the interest of U.S. Cyber Command in increasing its numbers by nearly five-fold, from the current thousand or so, over the coming years.

Dorothy Denning and Bradley Strawser also work with the airpower analogy; but in their case they focus specifically on the matter of the ethical appropriateness of various defensive actions. Their typology of defensive options—a spectrum of measures ranging from active to passive, and internal to external—is both comprehensive and illuminating, and the parallels between air and cyber defense are many. Interestingly, there is one major difference: active-external air defense measures run the ethical risk of causing destructive collateral damage, whereas similar cyberspace-based defensive actions are likely only to disrupt in unintended ways, not destroy. Thus, from their point of view, there may be more freedom to act—perhaps even to pre-delegate—in the cyber realm than in the domain of traditional air defense. But disruption may prove extremely costly, especially in an increasingly interconnected world. This is a telling point made by the contribution that explores the economic warfare metaphor.

ECONOMIC DISRUPTION AND COLLAPSE

British attraction to and use of economic warfare during the opening months of World War I a century ago—what we might call a *Britskrieg*—raises the intriguing parallel notion of state actions in cyberspace destabilizing the global economy today. The British put their plan into effect swiftly, but aborted the operation within weeks—probably the only strategic plan that was ever called off because it was too successful. The

collateral damage was immense. The Great Crash of 1929 was terrible and costly, but the British effort in 1914, if not swiftly curtailed, threatened to have global effects that would have inflicted far greater damage than that suffered fifteen years later. Nicholas Lambert tells the tale of how strategic economic attack on an early version of a global economy had immediate material impacts on banking and trade, and also on human psychology. Thus the great power, Britain, whose naval mastery ensured control of trade routes, and whose roles in global finance and communications were central, found itself in a position wherein massive disruption wrought unacceptable levels of collateral damage on friend and foe alike—and on British interests perhaps most of all.

The similarity with strategic cyber warfare today is eerie. One cannot target with precision because of the complexity of the world economic system. Costly, unintended impacts on neutrals may affect a state's decision to launch a cyberwar at economic targets in the first place. Further, one's own commercial sector might prove highly resistant to being party to such a campaign. At the outset of World War I, it seemed that there was no such thing as a blindly patriotic businessman, and the cooperation of the multinational companies of the day was hardly a given. British financiers refused to cooperate with their government in 1914 because they had little confidence that the latter understood or could deal with the costly economic disruption and dislocation that would ensue. In the cyber realm today, where the government does not control everything—perhaps not even very much—what level of confidence can there be that private companies will cooperate as civil and military leaders would like them to during a cyberwar? The short, unhappy life of the Britskrieg offers a profoundly cautionary tale, one that suggests the value of analogies may in some cases lie in learning to avoid rather than to imitate them.

A CULTURE OF INNOVATION

Cyber aspects of the Information Revolution portend a full, combined-systems transformation—one that hearkens back to similarly profound shifts, such as those that accompanied carrier aviation and amphibious warfare over seventy years ago, and AirLand Battle doctrine and battle management systems in more recent decades. Many involved in the business of strategic thought know that transformation is afoot once again, and that there are many factors that will influence the course of events—a course that is never predetermined, linear, or without setbacks. In past instances, organizational innovation was led by senior officers with traditional credentials who reacted not only to intelligence about specific potential enemies, but also to structural changes in the overall external threat environment. They were often, but not always, willing to experiment with fresh ideas, develop innovative doctrines, and create new promotion pathways for junior officers who showed promise as practitioners of the emerging modes of war. But today military innovations can hardly be the sole province of the armed services.

As John Kao points out in his study of the innovation phenomenon, a much more holistic approach is now necessary. Today there are over fifty countries that have broad, ambitious innovation and stewardship programs. And many of these technological innovations are already leading, or soon will lead, to new capabilities for waging war. If there is one part of the world that is about innovation and trying to avoid doctrinal lock-in as much as possible, Kao observes, it is Silicon Valley. A key factor in Silicon Valley's success is its permissive environment for linking academic research and business. For example, successful commercial spin-offs from research originating at Stanford University abound. Other important factors include the wide acceptance of social norms of trust, as well as openness to experimentation and risk-taking. Ultimately, innovation thrives when there is a galvanizing narrative around which to create a call to action and a call to arms. A compelling story is likely to be a "great attractor" of talent. The task now, as Kao notes, is to craft a narrative for cyber security—e.g., perhaps tied to the notion that individual, commercial, and national security all depend upon it—that will mobilize a wide-ranging, societal-level response.

THEORIES, TYPOLOGIES, AND ANALOGIES

The final two entries in this anthology serve, to some extent, as "conceptual bookends." Keir Lieber skillfully dissects the classic international security theory of the offense-defense balance, long thought to hold the key to understanding whether the world will be characterized by much conflict (when offense is easy) or

stabilized through robust deterrence (when defense is dominant in military affairs). He provides fresh perspectives that challenge this formulation, but goes even deeper, disentangling tactical military considerations of offense and defense on the battlefield from the larger strategic decision-making process about whether to go to war—or not. His insights are particularly valuable for the current discourse on cyberwar, which today is characterized by a general consensus that it is much easier to mount cyber attacks than to defend against them.

While this may be true, it is less clear that an offensive advantage in cyberwar will make conflict any more likely than if defense were perceived to have the advantage. The culprit here is the problem of uncertainty about effects. There is a very significant degree of unpredictability any actor who would launch a cyber attack must face—e.g., even the very specifically targeted Stuxnet worm "escaped into the wild" from the Iranian nuclear facility in which it was supposed to remain until it deleted itself. Now a range of malefactors are almost certainly reverse engineering and weaponizing modified versions of Stuxnet. Unpredictability of this sort may seriously undermine the logic about how offensive advantage leads to more war—a point that echoes Lambert's insights into the initial British enthusiasm for waging economic warfare and the swift onset of their regret at having employed this approach. The issue of predictability *per se* aside, an overarching issue here remains—indeed, the jury is still out—whether uncertainty about effects will lead to stability or conflict in the cyber realm. Nonetheless, the governing fact for cyber security today is that defense is not only more difficult than offense; it also requires a lot more resources. Under what conditions are governments going to be willing to make substantial, sustained investments in defense?

Perhaps one of the ways to proceed is to explore alternative defensive strategies that might prove less costly than the current firewall-and-antiviral paradigm that dominates the cyber security field today. With this in mind, Robert Axelrod concludes the anthology with a dazzling *tour d'horizon* of analogies drawn from history, biology, trade relations—even market-based commercial insurance underwriting standards. There is far more than Pearl Harbor in the historical "bag" of metaphors, Axelrod notes. He finds, for example, behavior-based notions of arms control—like the Chemical Weapons Convention—to be very much worth contemplating. From the life sciences he identifies analogies in the realms of biodiversity and herd immunity. And then he flips the (seemingly unusable) notion of economic warfare on its head, using the model of the insurance industry as a way of encouraging, in an actuarially-based fashion, some major improvements in cyber security.

Not all of his analogies are about limiting cyberwar; many offer thoughts about ways of taking the offensive. We are particularly drawn to his exploration of the potential for highly effective doctrinal innovations, such as were demonstrated by German airborne troops in the glider-borne assault on Eben Emael, and by the U-boats with their wolfpack "swarms." Axelrod makes clear that unexpected new tactics can have profound effects. The Chinese analogy that he considers at some length is also quite interesting—as Beijing seems attuned to sending strong signals of intent in the middle of a crisis. Cyber might provide a perfect means for signaling an adversary—short of outright acts of war. Overall, Axelrod's chapter concludes the anthology by offering up suggestions for the pursuit of several new arcs of research aimed at mining analogies and metaphors for fresh insights.

THE CASE FOR ANALOGIES

All of us on the cyber analogies team hope that this anthology will resonate with readers whose duties call for them to set strategies to protect the virtual domain and determine the policies that govern it. Our belief is that learning is most effective when the concepts under consideration can be aligned with already-existing understanding or knowledge. People use analogies, metaphors, and parables, both explicitly and implicitly, to link what is new to what is already known, as a bridge between the familiar and the new. Analogies at their best facilitate communication, particularly when carefully chosen to resonate with an identified audience. Since cyber security is a grand challenge that requires a whole-of-government approach, in close partnership with the private sector, analogies can be useful in communicating with diverse stakeholders in terms aligned with their present understanding, knowledge, and experiences. Analogies can also be a vehicle to engage active operators in useful discussions with informed people who are "outside the bubble" in which they perform their duties.

Cyber issues are inherently tough to explain in layman's terms. The future is always open and undetermined, and the number of actors and the complexity of their relations are too great to give definitive guidance about future developments. Historical analogies, carefully developed and properly applied, can help indicate a direction for action by reducing complexity and making the future at least cognitively manageable. There is a rich trove of analogies from which to choose, as this anthology demonstrates. And this is a good thing, given U.S. Cyber Command's serious interest in the systematic use of analogical reasoning as a means of gaining strategic insights and generating alternative options and courses of action. Above all, though, the chapters that follow should be viewed, and judged, in terms not only of their ability to stimulate the discourse on cyber security, but also by their contribution to winning H.G. Wells's "race between education and catastrophe."

Emily O. Goldman & John Arquilla

Fort Meade and Monterey

February 2014

The Cyber Pearl Harbor

James J. Wirtz

The notion that the United States is vulnerable to a strategic surprise attack bent on incapacitating computational and communication capabilities, which is often characterized by senior officials, military commanders, scholars and the popular media as a "Cyber Pearl Harbor," is a mainstay of the current strategic discourse.[1] The notion is reinforced by recurring media reports of ongoing, sophisticated and apparently successful foreign and domestic efforts to penetrate and disrupt military, commercial and private data bases and networks with an idea of compromising information, gaining clandestine control of various commercial and military systems or limiting or destroying the ability to access the World-Wide-Web or classified computer networks. Applying the "Pearl Harbor" analogy to characterize a cyber attack conjures up compelling images of a "bolt-from-the-blue" surprise attack in American political and strategic culture. It suggests that policymakers and intelligence officials will fail to anticipate the scope, nature or target of a cyber attack and that the U.S. military or technological infrastructure will suffer catastrophic paralysis, rendering them unable to develop a militarily or politically effective response in wartime. The analogy is also a clarion call to action: appropriate steps must be undertaken immediately in order to avert a potential catastrophe.

A more sophisticated and historically informed retelling of the events leading up to the December 7, 1941 attack on Oahu would suggest that the Pearl Harbor analogy is somewhat misleading. The attack occurred following a prolonged and serious deterioration in the relations between the United States and Japan. Additionally, the possibility of a carrier strike on U.S. military facilities on Oahu was anticipated and assessed in war games undertaken by the U.S. Navy. Indeed, protests by senior admirals over the forward deployment of the bulk of the U.S. Pacific Fleet to Hawaii were delivered to the highest levels of the Franklin D. Roosevelt administration.[2] Navy officers believed that the movement of the U.S. Pacific Fleet was a hollow gesture that would do little to bolster deterrence while leaving the force vulnerable to enemy action. In other words, the Japanese strike on Pearl Harbor can be characterized as something other than a bolt-from-the-blue attack. Nevertheless, using the Pearl Harbor analogy to characterize a potential cyber attack does provide insights into a specific political, strategic and operational setting that makes strategic surprise, deterrence failure and war more likely. The United States is vulnerable to a "Cyber Pearl Harbor" and it is possible to anticipate why and how it will occur.

To explain this assessment, this paper first describes the strategic setting that makes the United States a potential target of a surprise attack by describing how this strategic context shapes the incentives and perceptions of the parties in a potential conflict. It then provides a brief description of the operational setting that makes cyber attack an attractive element of a strategic surprise on the United States. The paper also discusses the serious challenges facing intelligence analysts as they contemplate the prospect of a Cyber Pearl Harbor, and how the analogy offers some important insights into analytical challenges involved in offering specific event prediction of cyber attack. The paper concludes with a suggestion of one possible way to reduce the likelihood of a cyber-based surprise attack on the United States.

THE STRATEGIC SETTING

Although operational or tactical surprise in war is universally endorsed as a force multiplier, relying on surprise as a "war winning" initiative at the outset of hostilities is an extraordinarily risky gamble. It is attractive to the weaker party in a potential conflict because of its military inferiority vis-à-vis the stronger party. One of the most common misconceptions surrounding surprise attack is that it occurs because the weaker party launching the attack overestimates its military prowess; in fact, both the weaker and the stronger party, for that matter, generally possess an accurate picture of the military balance. Because the weaker party recognizes its military inferiority, it seeks to develop various stratagems to circumvent a stronger opponent's military might in order to achieve some *fait accompli* or to alter the incentives of the stronger

party to reduce the likelihood that it will act on its deterrent threats.[3] A strategic surprise attack is one of those stratagems because it suspends temporarily the dialectic of war by removing an active opponent from the battlefield. Strategic surprise attack allows the weaker opponent to achieve objectives that it realistically could not expect to secure if it faced a militarily superior and fully engaged opponent.

If the weaker party is attracted to strategic surprise, the stronger party finds it difficult to detect and respond to indications of what is about to unfold. It too has an accurate assessment of the military balance, which suggests that deterrent threats are strong, credible and can be readily executed against a weaker opponent to good effect. Evidence of impending attack is often dismissed as incredible *ex ante* by analysts, commanders and officials because in their minds such initiatives are doomed to failure at the hands of their vastly superior military capabilities. The stronger party's decision-makers assess the potential actions of their weaker opponents with an "attritional mindset": they view a potential conflict as a "force-on-force" affair that the weaker party cannot realistically hope to win. Ironically, the weaker party also shares this perception, which leads it to seek ways—e.g., strategic surprise—to achieve its objectives by circumventing or paralyzing the military might of a vastly stronger opponent.[4]

strategic surprise attack temporarily suspends the dialectic of war

The strategic setting thus shapes the perceptions of the stronger and the weaker party in ways that make a strategic surprise likely to occur. The weaker party becomes extremely risk acceptant because it sees an opportunity to obtain objectives that have been deemed to be impossible to achieve in the face of a vastly superior military opponent. The stronger party will find it hard to respond effectively to indications of what is about to unfold because they will be appear too farfetched or "hare-brained" to be taken seriously. Even more to the point is that recognition that an attack would force decision-makers within the stronger party to accept that their military might and deterrence strategies, developed at great cost and effort, are about to fail to prevent war. Large bureaucracies are unlikely to respond effectively to such estimates because they appear to be *prima facie* irrational and threaten the personal, professional, institutional and political interests of those charged with responding to warning.

Given the current strategic setting, it is fairly safe to assume that the United States is more likely to be the victim, not the initiator, of a Cyber Pearl Harbor. As the militarily dominant state in the world today, it faces several potential opponents that are militarily inferior and who cannot realistically expect to achieve their objectives in the face of concerted and coordinated military opposition on the part of the United States and its allies. The extremely risky gambit of strategic surprise will appear attractive to these opponents as a means to sidestep superior military capabilities to present the United States with a *fait accompli* or to alter the strategic setting to circumvent existing deterrence strategies and to alter the incentives faced by U.S. decision-makers to engage in attritional combat to return to the status quo. The use of the Pearl Harbor analogy to explain the strategic origins of a cyber surprise attack captures the strategic and psychological dimensions of a brewing conflict that could lead the Unites States to fall victim to a strategic surprise attack.

THE OPERATIONAL DIMENSION

For decades, the U.S. military has led its competitors in exploiting the ongoing Information Revolution as a force enabler and has incorporated it as a key component of virtually every facet of its operations, infrastructure and doctrine. Military interest in the Information Revolution emerged as a solution to a long-forgotten strategic problem confronting the United States and its allies in the 1970s. It was feared that the North Atlantic Treaty Organization (NATO) would lose a mobilization race against the Warsaw Pact in a conflict along the inter-German border: NATO could match the mobilization pace of Warsaw Pact first and second echelon units, but would eventually be overwhelmed by third echelon units that could reach the battlefield faster than allied reserves. To stop this third echelon from reaching the forward battle area, the United States and NATO undertook several initiatives—"Emerging Technologies," "Follow-on-Forces Attack," and AirLand Battle Doctrine—to engage logistical bases and staging areas deep behind Warsaw Pact lines. These initiatives incorporated emerging intelligence, surveillance and reconnaissance

capabilities, information management technologies and long-range, precision-strike capabilities to penetrate Warsaw Pact defenses to disrupt the ability of the Soviets and their allies to maintain the tempo of forward operations and to prevent reserves from reaching the battlefield. By the 1980s, these initiatives sparked a lively debate about the nature and existence of a so-called "Revolution in Military Affairs" (RMA) and about whether the Information Revolution would fundamentally transform conventional warfare.[5]

Today, debate continues regarding the impact of the Information Revolution on war; but it is also clear that the United States still leads the world in the exploitation of computer and communication technologies as the basis of virtually all facets of its military operations.[6] The general impact on force structure, operations, and logistics of the Information Revolution also is relatively clear. The information revolution, combined with superior space and aerial surveillance capabilities, provides commanders with an exquisite ability to survey entire theaters of war in real time, identifying troop concentrations, command centers and critical communication networks for destruction by long-range, precision-strike capabilities, which deny an opponent the opportunity to engage in operational or even tactical maneuvers or the ability to undertake combined-arms operations. Once paralyzed, the opponent is intended to be rapidly overtaken by combined arms operations launched by highly professional and well-equipped, if numerically inferior, U.S. forces. Operational forces are also supported by a minimal logistical infrastructure as "just-in-time" supply chains deliver materials to forward units quickly and efficiently. In terms of air and conventional ground actions, this philosophy, technology, and doctrine has given the U.S. unmatched military capabilities in air and conventional ground operations, although the effects of this "system-of systems" approach have yet to manifest convincingly in small-unit or unconventional operations (e.g., counter-insurgency) and remain relatively unknown in war at sea.

No one disputes the fact that the U.S. military relies heavily on the Information Revolution as a basis of its military superiority; yet, there is limited acknowledgement of the fact that this reliance provides potential asymmetric advantages to opponents that cannot hope to defeat a fully engaged and information enabled U.S. military. Opponents have few viable alternatives to engage the United States effectively—superior numbers, firepower or a technically superior weapon (e.g., an advanced battle tank) can be rendered ineffective if they cannot be incorporated into operational and tactical maneuvers or combined arms operations. Nuclear attacks remain as a viable alternative to U.S. conventional superiority, but such attacks risk altering the political and strategic setting in an adverse manner for the attacker, which could lead to quick catastrophic defeat if the United States chooses to forgo conventional operations by relying on its own nuclear capabilities to achieve its objectives.

When compared to the drastic political and strategic consequences of even a limited nuclear attack on U.S. or allied forces, cyber attacks hold out greater promise of success. Cyber attacks can strike at the linchpin of U.S. forces, by rendering them at least temporarily unable to execute their preferred doctrine and operations, leaving them outgunned and outnumbered on some battlefield. The effects of cyber attack might be localized, and their impact relatively fleeting, but they do hold out the promise of at least temporarily removing the United States as an active opponent in some theater of operations. Moreover, because they hold out the possibility of removing the United States as an active opponent while inflicting minimal casualties or damage to U.S. forces, they can alter the political and strategic context of a confrontation by presenting U.S. policy-makers with a *fait accompli* at minimal cost to both parties. The burden of escalation would then shift to U.S. policymakers, who would have to choose war over political compromise. In other words, the degradation of U.S. military capabilities caused by a cyber attack will eventually be overcome, but the political

and strategic consequences of a failure of deterrence will linger long after U.S. forces recover their full capabilities. The history of strategic surprise attacks suggests that opponents often bank on these altered political and strategic circumstances to guarantee the eventual success of their risky gambit. They understand that even a wonderfully successful strategic surprise attack will not disable a vastly superior opponent forever; but they do hope to cause the opponent to come to believe that the game is not worth the candle. Needless to say, basing an operation on the expectation that a stronger opponent will choose not to respond to an attack is not in the best traditions of diplomacy or the military profession, which is an observation that further makes such scenarios unlikely to be taken seriously by the stronger party to a potential conflict.

In an operational sense, a cyber-based surprise attack on the United States would be attractive to a weaker opponent because it would offer a potential way to temporarily remove the United States as an active opponent on the battlefield. Furthermore, by allowing the opponent to achieve objectives quickly and with virtually no opposition, it creates the possibility of confronting the United States with a *fait accompli* achieved with minimal death and destruction. This would make this operational solution to U.S. military superiority politically and strategically attractive because it shifts the onus of escalation onto U.S. policymakers. Confronted by a deterrence failure, U.S. policymakers would be forced either to reassess their objectives or to engage in an inevitably more costly war to reverse initial losses. Unlike a nuclear surprise attack, which entails enormous destruction and political costs at the outset, a cyber-enabled surprise attack might just make living with a *fait accompli* a less costly alternative to war in the minds of the victims of a Cyber Pearl Harbor. A cyber attack, combined with conventional operations to achieve a *fait accompli* on the ground, provides the weaker party with a politically attractive and a potentially operationally effective means to temporarily prevent the information-empowered U.S. military from making its presence felt on some distant battlefield.

THE INTELLIGENCE CONTEXT

The Cyber Pearl Harbor poses several challenges to the intelligence community and to the officers and policymakers who have to take effective action in response to indications and warning of an impending attack. The first issue concerns the need to gain a general acceptance of U.S. military vulnerability to a cyber surprise attack. Acceptance of a potential vulnerability is based on a net assessment produced by the Intelligence Community or those charged with cyber operations that flies in the face of service cultures and the perception among "war fighters" that combat capability is primarily derived from traditional operations and weapons systems. By contrast, warnings about a Cyber Pearl Harbor highlight the importance of a part of the U.S. military infrastructure that war fighters generally prefer not to think about at all—support elements and personnel that by definition will never be engaged in direct combat. In other words, there is a tendency to avoid the fact that the U.S. military relies on the synergies enabled by the Information Revolution to produce its superior combat power and to treat those who manage the "information elements" of the combined arms team as second-class citizens who are less important than the "war fighters." Recognition of the cyber threat would thus upset overall defense hierarchies and preferences, which would threaten everything from promotion patterns to budgetary decisions to bureaucratic pecking orders. Things were never quite the same for battleship admirals after Pearl Harbor, but it probably is impossible to undertake quickly effective bureaucratic reform or cultural change in the absence of the motivation generated by catastrophe. Intelligence assessments based on long-term projections of emerging cyber trends and risks pose an immediate threat to personal, professional, bureaucratic and cultural preferences and will not be welcome news to their most critical audience—senior officers who benefit from existing hierarchies and budgetary priorities. Recognition of the cyber threat means that strategic choices will have to pay less homage to organizational preferences.

In terms of operational or tactical warning of a cyber attack, the challenge is truly daunting. Estimates vary, and the specifics remain highly classified, but it would appear that indications and warning of an impending cyber attack will probably be received, if at all, somewhere between a few days to a few seconds before some nefarious event. As a result, the time available for warning will not correspond to "political time"—the time needed for officers or policymakers to assess

the threat and to match political preferences to operational or tactical choices. A response to a cyber attack will have to be based on pre-established procedures that are implemented not by humans, but by the hardware and software that constitutes the U.S. information infrastructure. Estimating the specific methods, targets, timing and impact of a discrete attack might also be beyond the realm of possibility, especially if indications of what is about to transpire precede an attack by only a few seconds. These time constraints would call for the adoption of an indications and warning methodology when it comes to responding to the threat of cyber attack, and the development of protocols and standard operating procedures to increase defenses in response to a heightened threat environment. To beat such stringent time constraints, intelligence analysts who focus on the cyber realm need to produce real-time threat assessments in order to enable defensive measures that can deter and, if need be, mitigate and defeat an attack.

Events surrounding Pearl Harbor, or even the 9/11 attacks, however, provide ready examples of responses to warning that failed to meet imminent threats. At Pearl Harbor, warnings of impending hostilities prompted actions to prevent sabotage, which actually left U.S. facilities more vulnerable and less responsive to air attack. In the hours prior to the 9/11 attacks, the would-be hijackers attracted suspicion, but the heightened security protocols they tripped proved ineffective when it came to disrupting their scheme. Intelligence analysts, and policymakers for that matter, have to ensure that the changes in security protocols and standard operating procedures that are enacted in response to warning actually are an effective response to imminent attack. This requires long-term analyses and net assessments to understand vulnerabilities and to devise responses that are likely to deal effectively with potential threats.

> In terms of operational or tactical warning of a cyber attack, the challenge is truly daunting.

A third problem facing intelligence analysts is the possibility of technological surprise that can be produced by strategic interactions, narrow technical vulnerabilities or even "normal accidents." Because a strategic surprise attack that incorporates a significant cyber dimension has never occurred, it is difficult to predict how extremely complex technical systems and human-machine interfaces will respond to novel events or unanticipated technical exploits.[7] An interaction between competing offenses or offensive and defensive programs might yield heretofore unknown consequences among information systems and specific weapons, information systems and conventional forces writ large, or information systems and critical support functions (e.g., logistics, electrical grids, etc.). Information infrastructures might also fail quickly and catastrophically if heretofore small and seemingly unimportant technical vulnerabilities can be exploited across entire networks or are used to destroy key nodes in important communication architectures.[8] Cascading failures also might occur because of operator error or because standard operating procedures actually exacerbate unanticipated degradation of networks. Because the information infrastructure permeates all facets of military operations, technological surprise becomes an acute problem because the behavior of emerging information systems is not completely understood, especially in circumstances that are not specifically anticipated by designers and operators. When it comes to technological surprise, cyber attack also holds out the possibility of producing systemic effects, which stand in contrast to the general tendency of technological surprise to produce effects related to specific weapons and localities.

It is probably too much to expect that intelligence analysts will be able to offer policymakers specific event predictions when it comes to a strategic surprise attack based on cyber operations. The information realm is not entirely understood, opportunities for technical exploitation by the opponent are not completely identified, and not enough time is available for policymakers or officers to mull over options to select appropriate efforts to deter or defend against an attack. Intelligence analysts might be able to issue general warnings of a heightened threat, but an appropriate response will still depend on sustained analysis that can generate appropriate responses to potential attack vectors. Confusion over responsibilities and unwarranted assumptions about the likely behavior of other services or commands and inappropriate defensive preparations lead the defenders at Pearl Harbor to take ineffective or lackadaisical responses to "war warnings," last minute indications of an approaching enemy, and even reports of an engagement with enemy forces. From an intelligence perspective, it

is easy to see how a cyber attack, coupled with conventional operations, could produce similar results.

THE WAY AHEAD

The specific details of the Cyber Pearl Harbor cannot be known in advance, but the general outline of the scenario can be described with some certainty by relying on the history of previous instances of strategic surprise attack. A militarily inferior opponent will utilize a cyber attack in conjunction with conventional military operations to present the United States with a *fait accompli*, probably at surprisingly minimal cost in death and destruction. This failure of deterrence will occur in an unanticipated manner, and might even be abetted by technological surprise—utilizing unrecognized technical vulnerabilities, undertaking heretofore unimagined types of operations, exploiting operating errors, triggering ineffective standard operating procedures, or causing catastrophic cascading effects. The cyber attack itself will be highly asymmetric in the sense that the attacks will be subtle and undertaken at relatively minimal cost, but they will incapacitate extraordinarily powerful forces developed and deployed at enormous expense. The effects of the Cyber Pearl Harbor will be temporally and geographically limited, but it will have a lasting political and strategic impact on U.S. interests. The Cyber Pearl Harbor will confront the United States with a decision to fight or to accept a significant political or military setback given deterrence failure and an altered strategic setting. The opponent will expect that the United States will not bring the full force of its superior military capability to bear to return to the *status quo antebellum*. The architects of the Japanese attack on Pearl Harbor, for example, informed their political leaders that they could give them free reign in the Pacific for six months, suggesting that a rapid diplomatic settlement was in their interests. Although that estimate turned out to be a bit optimistic, it might still serve as the operational time horizon for a future opponent that risks its fate on a Cyber Pearl Harbor.

operational or tactical solutions to cyber vulnerabilities are difficult to devise and implement

Policymakers and officers have demonstrated an increasing recognition of this scenario; their repeated references to a potential Cyber Pearl Harbor reflect a growing appreciation of the threat faced by the United States. What is missing is a ready solution to mitigate the threat. Technical solutions to the problem seem unavailable—as soon as technical vulnerabilities are identified and mitigated new vulnerabilities are discovered or are even created by "patches" that address known problems. Operational or tactical solutions to cyber vulnerabilities are difficult to devise and implement because of the inability to bridge professional, cultural and bureaucratic divisions within a complex defense establishment that privileges the interests of "warfighters" over other elements of an information-enabled military. The asymmetric threat posed by cyber attack is noted, but solutions to the problem are largely confined to technical matters that must be executed in a way that preserves bureaucratic and professional preferences. Intelligence is unlikely to produce specific event predictions in sufficient time for policymakers or officers to identify and select appropriate technical, tactical, or operational responses.

It thus might be suggested that the best way to deal with the possibility of a Cyber Pearl Harbor is to prevent it from happening in the first place. In other words, a strategic solution to the problem of bolstering deterrence might be the best way to prevent this scenario from unfolding. The exact deterrent strategy to be adopted, however, requires clarification. Deterrent strategies based on punishment (i.e., inflicting prompt damage out of proportion to the gains achieved through aggression) or retaliation (i.e., inflicting death and destruction until the opponent abandons their ill-gotten gains or alters unwanted behavior) will not be effective because those contemplating a strategic surprise attack have deliberately developed their plans with an eye towards nullifying these types of strategies. The attacker hopes to alter the political and strategic calculus of the victim of surprise in a way that reduces their incentives to act on previously stated deterrent threats based on punishment and retaliation. The attacker generally recognizes that the victim of surprise will retain the ability to act on deterrent threats; instead, they hope that acting on deterrent threats will appear politically and strategically unattractive once deterrence fails. This turn of events might appear incredible to the potential victim of surprise *ex ante*—a perception that sets the stage for surprise to occur—but it is often the weak reed that is

grasped as a theory of victory by those who gamble everything on surprise. Thus, restating or reinforcing deterrent threats based on retaliation or punishment will do little to prevent The Cyber Pearl Harbor. Ironically, the Roosevelt administration moved the Pacific Fleet to Pearl Harbor to bolster deterrence, but in so doing actually weakened deterrence by placing its retaliatory force within range of a pre-emptive strike. The party that initiates surprise attack already recognizes that it will be lambasted if the stronger party brings its full military capability to bear. Messages that restate the military superiority of the stronger party in a potential conflict make little impression on a weaker party that has already embraced the gambit of strategic surprise.

Instead of deterrence by retaliation or punishment, deterrence by denial offers the more promising path to prevent the occurrence of The Cyber Pearl Harbor. Denial implies a recognition of the threat of cyber surprise attack and the adoption of pro-active program to decrease an opponent's confidence that a cyber attack will actually succeed. Clearly, a technological response to the threat of cyber attack is in order, but tactical, operational, and strategic programs to mitigate the loss of information-enabled systems might decrease at least the appearance that cyber attacks can have a systemic impact across the entire U.S. defense establishment. Several general ways to bolster deterrence by denial can be suggested. Information-based systems, for instance, might be more closely integrated in an organic fashion in operational units, decreasing their reliance on a single operating network. Wargames and experiments might be conducted to test and operate systems or combat units in a situation where information systems are degraded in an effort to design wartime protocols or work-arounds. Wartime systems might be developed so that they can be constituted in times of national emergency to replace systems and networks that have been compromised by cyber attack. Most importantly, information systems need to evolve continuously to prevent them from becoming a static target that can be exploited by long-term efforts by potential opponents. By creating an ever-changing operational picture, potential opponents might never have confidence that they have managed to penetrate information systems in a way that may produce strategic damage against U.S. warfighting capabilities. Static technology, doctrine and operations give opponents the time to discover technological or operational weaknesses that might form the basis of the conclusion that it just might be possible to paralyze U.S. forces long enough to create a *fait accompli*.

Creation of a dynamic deterrence-by-denial capability is no small matter. It would be expensive and redundant. It also lacks clear-cut metrics related to sufficiency, other than a clear indication of when insufficient effort has been expended and deterrence has failed. A dynamic deterrence-by-denial program also flies in the face of economic logic and the philosophy that animates the Information Revolution. Instead of a single, highly-connected system, a robust denial capability might require high redundancy or systems that can retain important functionality without networking. Information systems work best when everyone is free to communicate with everyone else, but such systems could be vulnerable to cascading failure if damage to some seemingly unimportant parts of the systems propagates across the entire information infrastructure. It would also be prudent to search for vulnerabilities across the interface between information infrastructure and operating units. In other words, can single-point logistical, communication or command failures threaten entire theaters of operation? Have key systems and combat forces become static and vulnerable to penetration and potential disruption? Are otherwise robust operational capabilities overly reliant on vulnerable information-based networks for supply or command and control functions, creating critical weaknesses that will reveal themselves in wartime? Do the opponents really have to

act swiftly, or do they possess the luxury of time to seek out weaknesses and devise exquisite operational means to exploit their opponent's obvious military superiority?

A dynamic deterrence-by-denial effort cannot eliminate all of the vulnerabilities that are inherent to an information-based military, but it can reduce an opponent's confidence that a solution to their military inferiority can be found in an asymmetric attack. The key is to deny the opponent the confidence that some ingenious, albeit extraordinarily risky, plan might actually succeed in gaining an immediate *fait accompli*. For the stronger party, this effort must occur in peacetime. In war, the opening move is granted to the side launching a strategic surprise attack.

the best way to deal with the threat of a Cyber Pearl Harbor is to alter the incentives and perceptions of the opponent

CONCLUSION

If warfighters are reluctant to credit the synergies created by an information-enabled military as the linchpin of their combat power, then technicians charged with maintaining information capabilities are probably equally reluctant to embrace a strategic-based solution to the problem of cyber surprise attack. For them, technical solutions will beckon as the most effective and appropriate response to cyber attack. Nevertheless, if technical solutions are not informed by an awareness of the political, strategic, intelligence and operational context of surprise attack, they are unlikely to address the incentives faced by a weaker party as they consider ways to circumvent the warfighting capabilities and deterrent strategies of a stronger opponent. Deterrence by denial requires that the opponent perceives that there is little opportunity to achieve asymmetric effects by employing cyber attacks. The best way to deal with the threat of a Cyber Pearl Harbor is to alter the incentives and perceptions of the opponent so that they fail to gain confidence in a stratagem to achieve their objectives by sidestepping the superior military power of their opponent.

This paper also highlights the fact that, to be significant, a cyber surprise attack should not occur in a political vacuum. In the scenario outlined here, cyber attacks are coupled with conventional operations to achieve specific political or strategic objectives. To date, this sort of cyber event has not occurred in world politics, which suggests that official warnings about a Cyber Pearl Harbor do not describe ongoing cyber activities, but are rather a warning about the possible shape of things to come. ❊

NOTES

1 Keith B. Alexander, Memorandum for Record, Subject: United States Cyber Command (USCYBERCOM) Commander's Strategic Assessment for Operating in Cyberspace—Preventing a Pearl Harbor Environment, United States Cyber Command, Fort George G. Meade, Maryland, 23 March 2012.

2 Navy officers called for a return of the Pacific fleet to the West Coast to prepare for a potential war with Japan. They believed that the lack of adequate facilities at Pearl Harbor made forward deployment a hollow deterrent that only served to give the American public a false sense of confidence in U.S. defenses in the Pacific. When the Commander of the U.S. Navy in the Pacific, Admiral J.O. Richardson, failed to convince his superiors to re-position the fleet in California, he penned a message to President Franklin D. Roosevelt that led to his relief: "The senior officers of the Navy do not have the trust and confidence in the civilian leadership of this country that is essential for a successful prosecution of a war in the Pacific," see George W. Baer, *One Hundred Years of Sea Power: The U.S. Navy, 1890-1990* (Stanford: Stanford University Press, 1994), p. 151.

3 James J. Wirtz, "Deterring the Weak: Problems and Prospects," *Proliferation Papers*, No. 43, Fall 2012.

4 For a full description of how the strategic setting shapes the perceptions of the strong and weak parties in a looming conflict see James J. Wirtz, "Theory of Surprise," in Richard K. Betts and Thomas G. Mahnken (eds.) *Paradoxes of Strategic Intelligence* (London: Frank Cass, 2003), pp.101-116.

5 Dima Adamsky, *The Culture of Military Innovation: The Impact of Cultural Factors on the Revolution in Military Affairs in Russia, the U.S. and Israel* (Stanford: Stanford University Press, 2010).

6 Barry D. Watts, *The Maturing Revolution in Military Affairs,* Center for Strategic and Budgetary Assessments, 2011.

7 Charles Perrow, *Normal Accidents: Living with High-Risk Technologies* (Princeton: Princeton University Press, 1999); and Scott Sagan, *The Limits of Safety* (Princeton: Princeton University Press, 1993).

8 Ted G. Lewis, *Bak's Sand Pile: Strategies for a Catastrophic World* (Williams, California: Agile Press, 2011).

Applying the Historical Lessons of Surprise Attack to the Cyber Domain: The Example of the United Kingdom

Dr Michael S. Goodman[1]

Surprise is an integral component of all confrontations, political as well as military. As the mode of conflict has evolved, so too have the methods and tactics needed to mount surprise attacks; yet the essential variables—time, place, ground, direction, numbers and capabilities—have remained relatively constant. There is a large literature on 'surprise attack', and the lessons are overwhelmingly drawn from examples of the Twentieth Century. These generally involve instances of state conflict and traditional conceptions of 'war'. The Cold War arms race introduced some revisions, including the issue of 'technological surprise', but fundamentally the issues remained the same. At the center of the discussion has been an examination of 'intentions' and 'capabilities'. As the nature of warfare has changed, these two constituent elements have remained intact; indeed, although there are subtleties to them, they remain perfectly valid for a consideration of a terrorist attack or any other sort of asymmetric warfare, including cyber attack. Yet, cyber attack has its own unique characteristics that require a recasting of how 'surprise' is understood. This chapter considers some of the lessons that can be drawn from historical studies of surprise and focuses on their applicability to the cyber domain. In concluding that there is no guarantee to preventing surprise, it then considers how the United Kingdom approaches the problem of cyber attack.[2]

DEFINITIONS

There is no question that in warfare, surprise is a powerful tactic. It has been described as a 'force multiplier' and, perhaps as a result, is frequently employed as an asymmetric ploy. In one of the earliest but most frequently cited studies, Richard Betts argued that 'surprise is almost superfluous if the attacker dwarfs the opponent … vital if opposing forces are evenly matched … shock is the only hope at all if the victim is stronger than the attacker'.[3] Invariably, it is the device of the weaker side against the stronger but generally does not result in the attacker (and weaker side) ultimately emerging victorious.[4] This should not confuse the aims of the initial surprise attack though; it might not have been intended to cause overall victory, being simply an attempt to ameliorate the present position. Reflecting its utility in conventional warfare, 'surprise' occurs because of a failure to provide or respond to relevant intelligence in three principal ways:

> **Strategic warning**—This is a long-term warning, often weeks or months in advance, usually following evidence of large-scale troop deployment. What is important here is that 'strategic' warning denotes evidence of capabilities; true intentions are potentially concealed.
>
> **Operational warning**—This is imminent warning, often received only days before a potential attack. It might reveal the starting movements of the actual attack.
>
> **Tactical warning**—This is the immediate stage in the hours before an attack or some *fait accompli*. The decision to go has been taken, initial efforts have taken place and the attack begins. Almost certainly by the time warning has been received, it is difficult to pre-empt or forestall an attack.[5]

'Warning', in these cases, has frequently been replaced with words like 'failure', 'intelligence', or 'surprise' in the literature on surprise attack.

The means needed to effect strategic surprise might well be different from those needed for operational or tactical surprise. The actual time involved in these definitions has also become compressed with advances in technology. Thus, for instance, tactical warning in a conventional sense might still encompass a matter of weeks; whereas, in the missile era it might be just a few hours. The importance here is what the context reveals for the response time. To allow for counter-measures, but also to provide warning, reliably, of an impending attack, operational

warning might be more important than tactical warning in the information age.

These ideas can also be considered in terms of the impact of surprise. The connection between warning and consequence is clear. Betts identifies a subtlety where the surprise is not actually in the attack itself, but in the consequences that result: 'the victim may not be surprised that war breaks out, but he is still surprised in military *effect*'.[6] Betts also writes that 'effective surprise changes the ratio of forces, making the peacetime balance almost instantaneously obsolete'.[7] While this might be true, at least in the immediate aftermath of the attack, history has shown that the victim (usually the militarily stronger party) often will regain the initiative in the conflict after the shock of surprise begins to fade.

In the context of the cyber domain, another important consideration is 'technological surprise'. Patrick Morgan has written how 'most strategic surprises do not result from a crucial new weapon used on an unsuspecting opponent with devastating effect, i.e., the atom bomb. What we do find in many cases is the use of existing weapons and forces '*in new and different ways*'.[8] 'Technological' surprise changes the analysts' calculations somewhat—here the issue is as much about developing an accurate estimate of the opponents' capabilities as it is estimating their intentions. A distinction might be drawn in this context between war and peace: it might be assumed that in times of war a weapon will be deployed once it is ready; during peacetime just because a weapon is operational it might not (as in the case of nuclear weapons) be readily employed.

In contrast to Morgan's statement, the 'Defense Warning Office' of the U.S. Defense Intelligence Agency (DIA) has identified four types of technological surprise:

Type 1: A major technological breakthrough in science or engineering.

Type 2: A revelation of secret progress by a second party which may have an unanticipated impact.

Type 3: Temporal surprise, when a party makes more rapid development or advancement in a particular technology than anticipated … this type of surprise is often facilitated by technology transfer that accelerates progress beyond a traditional linear development cycle.

Type 4: Innovative technology applications, such as using an airplane as a weapon on September 11 2001 … such innovations do not necessarily require technical expertise, but rather the creativity to use available resources in a new way.[9]

What this classification of technological surprise suggests is the centrality of intelligence and a careful understanding of what is meant by 'intent' and 'capability' when thinking about the future actions of some opponent. Estimating intentions is a far less tangible and more speculative process than estimating capabilities.[10] It is important to distinguish between different types of intention: it is possible to talk in terms of a *strategic* intention—the desire to acquire a particular weapons capability in the first place—and a *tactical* or *operational* intention—the desire to actually use the weapon once acquired.[11] While these may be related, the former does not necessarily imply the latter. Capabilities, on the other hand, are generally more visible than intentions. They are physical manifestations which, depending on the technology, are difficult to conceal, especially because they are accompanied by a significant industrial or operational infrastructure if they are to be used on a large scale. 'Capability' can also be further defined because a *theoretical* or *latent* capability is a very different prospect than a *practical* capability, yet the two do not necessarily follow sequentially. The close

relationship between 'intent' and 'capability' and the blurring of their use can, and has, led to poor quality intelligence estimates. In other words, when an intention is assumed it can contaminate estimates of capabilities, or vice versa.[12]

From an intelligence perspective, there are four scenarios that highlight the complexities involved in producing accurate estimates. These cases demonstrate the difficulties in assessing technological advances and, in doing so, conflate issues of intent and capability. These are particularly instructive for the cyber domain, given the prevalence of technological innovation:

Case 1) We develop a cyber capability—They develop a cyber capability
This is the most frequent scenario and, as the title suggests, occurs when both sides develop comparable technologies. Although both parties may seek similar capabilities, they may take completely different routes to achieve them. It may thus become difficult to estimate which side in fact has superiority in terms of cyber capabilities.

Case 2) We develop a cyber capability—They don't develop a cyber capability
The challenge here is to verify that a potential opponent actually lacks strategic or operational capability to undertake cyber activity. An opponent might decide to hide the existence of a capability so that it can be used to maximum effect in a significant attack.

Case 3) We don't develop a cyber capability—They develop a cyber capability
In a sense, this is the most troublesome case. It can emerge from a political or strategic decision not to develop certain military capabilities, or because of a failure to enjoy or understand important, theoretical, technological or operational innovations in the cyber realm.

Case 4) We don't develop a cyber capability—they don't develop a cyber capability
This has similarities to the other cases, but is still unique. Once more the problem of proving a negative is evident—that a capability is *not* being built—as is the difficulty of assessing information about which no direct experience exists.[13]

In all of the situations described above, the word 'capability' can denote both defensive and offensive means. An examination of the preceding factors suggests different ways of calculating what constitutes a 'threat':

HIGH THREAT = high likelihood + high impact
LIMITED THREAT = low likelihood + high impact
LIMITED THREAT = high likelihood + low impact
LOW THREAT = low likelihood + low impact

This formula is instructive, though it is by no means infallible. Betts offers a useful statement that encompasses these issues: 'the most militarily telling innovations are those in which the development of a new possibility is coupled quickly with an appropriate strategic and tactical concept, and is applied promptly in battle before the enemy becomes aware of, absorbs, and adapts to it.'[14]

How can 'cyber attack' be characterised? It can be interpreted in many ways: by the originator (state-based, criminal, terrorist, individual etc.); by the method (hacking, denial of service etc.); by the aspiration (espionage, criminal etc.); and the effect (disabling a network, spreading propaganda, stealing information etc.). In short, like the term 'cyber warfare', it is a somewhat ambiguous phrase that has become a catchall term for almost any sort of electronic strike.

CAVEATING LESSONS FROM HISTORY

There are many books on intelligence failure and surprise attack. Although not the first, the field was defined by Roberta Wohlstetter's 1962 study on the attack on Pearl Harbor.[15] The majority of the detailed studies of surprise attack have focussed on conventional warfare. Even when discussion has extended to include terrorist attacks, the common links to orthodox military operations have been emphasised.[16] A few caveats need to be recognised before examining the lessons from history. The first is that the overwhelming majority of accounts focus on the reasons why the defending nation was surprised, not how the attacking side achieved it. For the purposes of

this paper that distinction does not pose a problem, but it is important to highlight that imbalance in the treatment of surprise and how the perspective can skew understanding. The second caveat is that many of the lessons are drawn from the same case-studies: Pearl Harbor, Barbarossa, the Korean War, the Cuban Missile Crisis and the Yom Kippur War are the most frequently cited and examined cases. As new evidence is declassified or becomes available, these cases have been subjected to ongoing reassessment. Very rarely is treatment of the cases based on complete access to all information. A third factor is related to this second caveat and raises the question of selectivity: surprise attacks are obvious but what about those instances where surprise was forestalled? By their nature they are harder to spot, but in fact it is just as important to learn the lessons from surprise prevention as it is surprise attack. Yet the literature is firmly tipped towards examples of failure.

Studies of surprise attack have tended to focus on different aspects of the phenomenon, often encompassing the elements of the intelligence cycle.[17] It has been thought that collection has been at fault, sometime analysis, and often policy reluctance or an inability to act. These are very simple conclusions though, and most of the work on the topic has attempted to delve much deeper. One of the most well-trodden paths is analysis, particularly the view that in every case there was some sort of warning which might not have been obvious at the time but which, nonetheless, was present. The conclusion that is frequently drawn is that if only it were known what to look for, then surprise could have been prevented. This logic has led to some very detailed analyses, often focusing on the sorts of psychological traps or barriers that can affect progress. In fact these are not limited to surprise but are everyday issues of concern to analysts; but the effect is magnified when the ramifications of various analytical pathologies are much more serious.

How might the system work in a perfect environment? Ideally the intelligence community will receive information that is reliable, verifiable, and time-sensitive. Policymakers will receive the intelligence assessments, will find them credible and will act. Crisis averted. But things are never this simple or straightforward; in the cyber domain, issues of timeliness make this particularly tricky. It is commonly assumed that cyber attacks happen at lightning speed—many certainly do—but the most destructive and effective ones are far more complex, involving significant lead-time and a long lifespan. What, then, is meant by 'warning'? David Omand devises a useful typology, distinguishing between:

Strategic notice: An estimate of what may happen in the future

Situational awareness: What is going on in terms of who, what, where, when etc.

Explanation: What is going on in terms of why, what for?

Prediction: Where next? What next? When next?[18]

In the cyber context, there is an important distinction to be drawn between early warning of an attack possibly coming and early detection of the attack itself. These imply that it is far too simple to draw lessons from the historical studies by concentrating on the individual elements of the intelligence cycle. The greatest lessons, therefore, emerge from more generic points:

the most effective cyber attacks are complex, involving significant lead-time and a long life-span

1. What the various studies seem to accept is that intelligence is absolutely integral to trying to prevent surprise. Historical surveys suggest that there is invariably a decent amount of intelligence, but often the specific questions needed to avoid surprise (what, where, when etc.) are absent.
2. It is important to remember that to label an attack as a 'surprise' is too simplistic. Frequently the decision to attack is not a surprise, but elements of the attack itself are. This suggests that it is possible to achieve tactical surprise without being surprised strategically.
3. Perhaps as a reflection of the analytic imbalance, it is usually found that surprise is a result of the defender's inabilities rather than the attacker's abilities: it is a question of one side's failure, not the other's success.
4. The asymmetric nature of surprise attack is often its most ignored feature: an examination of the strategic

> balance will usually be enough to convince the defender that their capabilities are sufficient to withstand an attack, which makes them reluctant to take drastic action to forestall or pre-empt what appears to be *ex ante* as a risky and improbable event.

The problem with historical analyses, as Klaus Knorr identifies, is that 'the utility of extracting lessons for statecraft from historical analysis rests on the assumption that past experience is relevant to the future.'[19] There are clear similarities that can be observed in case of conventional surprise attack, but how valid are these for the Twenty-First Century cyber domain?

IS THE CYBER DOMAIN SO VERY DIFFERENT?

By what measure is it possible to compare conventional warfare and the cyber domain? There are a host of variables to consider and none are as simple as might appear at first glance:

Purpose of attack

It is possible to conceive of various types of 'attack', loosely defined: for instance, espionage, sabotage, crime, or destruction. These forms are common to different tactics, whether conventional, irregular or cyber. The differences between them lay in the way that they are conducted and their intended effect. The literature on surprise attack is focused on military strikes but there are other sorts of examples to choose from, including diplomatic surprise (for instance the Nazi-Soviet Pact); operational surprise (such as the penetration of an organization by a spy); or longer-term failures (the failure to anticipate the end of the Cold War is a good example). The role of intelligence will vary depending on the initiative undertaken by the attacker.

Identity of attacker

Who is the behind the attack? This is not simply a question of who pressed the button, but who possesses the technical expertise to enable it? Is the 'attacker' the authorizer of the attack, the facilitator, the source of the technical operation, the location of the server, etc. These may not be the same, and determining who is behind an attack is particularly difficult for non-state actors. In conventional warfare these questions are far simpler to ascertain: there have been no instances of significant surprise attack where it is not immediately clear who the perpetrator was. The identity is less clear in the context of terrorist attacks but, by and large, even if evidence is not forthcoming, a group will often claim responsibility.

Nature of attack

The evolution of warfare has changed the nature, scale, and operability of attacks. Central to these variables is technology. Missiles greatly increased the speed (and devastating effect) by which war could be launched, just as aircraft had a generation earlier. Asymmetric tactics and conventional attacks are often devastating and the method will, almost certainly, involve the physical destruction of something. Attacks in the cyber domain are different for two crucial reasons: the victim might not realize that they have actually been attacked, possibly even after the attack is finished; and the attack is not kinetic, though the consequences of cyber attack might cause a kinetic effect.[20] This latter point can also be perceived in terms of tactical vs. strategic strikes. It is perfectly conceivable that a military attack might threaten the strategic future of a country, but can the same be said for a cyberattack?

> the victim might not realize that they have been attacked, possibly even after the attack is finished

Identity of victim

For conventional warfare and terrorist attacks the victim was invariably the state (in terms of its people or inhabitants). Although the goal might have varied, usually it had something to do with weakening or undermining the position of the state. In the cyber domain this is not necessarily the case. It is much more difficult to target a specific group of people in the same way that a terrorist attack might, depending on what the criteria are (race, religion, nationality etc.). That said, it is easier to target specific institutions through their electronic footprints—there are myriad examples of individual corporations being selected for attack.

Option to respond

The option to respond has to be based on evidence that an attack has taken place, something that might not always be

obvious. Unlike military attacks it will not always be evident who the attacker is. Furthermore, there are the questions of scale and proportionality: what sort of cyber-attack, for instance, will necessitate a military response and, if so, what form should it take? At what stage might pre-emption be considered? During the Cold War various systems were developed to cope with such eventualities. The North Atlantic Treaty Organization (NATO), for instance, had its own Counter-Surprise system based on a series of alerts. Once triggered these would suggest that an attack was under way and would set in motion pre-planned responses. The system enabled the supreme military allied commander, SACEUR, to counter immediately even when time did not allow political consultation.[21]

SURPRISE ATTACK IN THE CYBER ERA

Given the historical examples of surprise attack and the differences between the cyber domain and other forms of warfare, what lessons emerge?

Detection and Attribution of a Surprise Cyber-Attack

Knowing that an attack is imminent or has taken place is, of course, crucial. Detection and attribution are therefore paramount—the 'digital blood trail', as one member of Interpol has described it.[22] Clark and Landau argue that 'solutions to the "attribution problem" lie outside the technical realm', that instead political, diplomatic and legal measures should resolve any doubts. It is an interesting view but difficult to substantiate.[23] More importantly they suggest that 'attribution' is important for four reasons and, in doing so, conflate it with 'prediction':

Before the Fact—Prevention or Degradation
During the Fact—Mitigation
After the Fact—Retribution
Ongoing: Attribution as a Part of Normal Activity

The advent of international law ensured that formal declarations ought to have been issued to confirm that two nations were at war. Today's cyber context is markedly different and this distinction is crucial. There have been many debates as to what, precisely, should be the 'red line'—the point at which a cyber-attack is deemed sufficient to warrant a response, military or otherwise. Given that most attacks will be conducted without prior notice, in the cyber domain it is extremely unlikely there will ever be a repeat of the beautifully worded and terribly polite way in which the United Kingdom declared war on Japan in 1941:

FOREIGN OFFICE, DECEMBER 8 (1941)
SIR,

ON THE EVENING OF 7TH DECEMBER HIS MAJESTY'S GOVERNMENT IN THE UNITED KINGDOM LEARNED THAT JAPANESE FORCES, WITHOUT PREVIOUS WARNING IN THE FORM OF A DECLARATION OF WAR OR OF AN ULTIMATUM WITH A CONDITIONAL DECLARATION OF WAR, HAD ATTEMPTED A LANDING ON THE COAST OF MALAYA AND BOMBED SINGAPORE AND HONG KONG.

IN VIEW OF THESE WANTON ACTS OF UNPROVOKED AGGRESSION COMMITTED IN FLAGRANT VIOLATION OF INTERNATIONAL LAW AND PARTICULARLY OF ARTICLE I OF THE THIRD HAGUE CONVENTION RELATIVE TO THE OPENING OF HOSTILITIES, TO WHICH BOTH JAPAN AND THE UNITED KINGDOM ARE PARTIES, HIS MAJESTY'S AMBASSADOR TO TOKYO HAS BEEN INSTRUCTED TO INFORM THE IMPERIAL JAPANESE GOVERNMENT THAT A STATE OF WAR EXISTS BETWEEN OUR TWO COUNTRIES.

I HAVE THE HONOUR TO BE, WITH HIGH CONSIDERATION, SIR,

YOUR OBEDIENT SERVANT

WINSTON CHURCHILL[24]

The point here is not the formal response by which war is declared, but the manner in which hostilities break out initially—a surprise attack. Matters are further complicated because the question of attribution is neither simple nor straightforward.[25] In addition, how should a response be judged: If the perpetrator is a non-state actor, can a state be held accountable? What levels of proof are needed to ascertain state involvement? Does

it make a difference where the attacker is physically based, particularly if nationality and allegiance are different? For all of these reasons and more, an attack in the cyber domain is different from conventional military operations.

Defending a Surprise Cyber Attack

The clear lesson from historical studies of surprise attack is that there is no way to ensure against surprise. In the aftermath of the failed assassination attempt by the IRA on Prime Minister Margaret Thatcher at a Brighton Hotel, a statement was issued that said 'Today we were unlucky, but remember we only have to be lucky once. You will have to be lucky always.'[26] Clearly the definition of 'failure' will differ depending on whether the attacker or defender's perspective is taken.

The most productive way of thinking about defending against a surprise cyber attack is to conceive it in either a strategic or tactical context. The lesson from the performance of Britain's most senior analytical body, the Joint Intelligence Committee (JIC), during the Cold War is that, while strategic warning was almost always given, tactical knowledge was often lacking. There are several issues that are raised by this observation. The first is to ask, reasonably, what should policymakers expect from their intelligence community and the JIC in particular? Is the provision of timely, strategic warning enough, especially when provided days or weeks beforehand? Having tactical intelligence on the 'decision to go' would, of course, be ideal, but regardless of whether this is possible, does it really make much of a difference, given that the response time might be reduced to minutes? This has clear implications for the cyber era, when time sensitivities are arguably far more crucial.

David and Sakurai write that 'at the tactical level of specific attacks, it is almost impossible to design systemic strategies for identifying the immediate threat details of exactly where, when and how an attack will occur. However, at the operational level, how cyber terrorists [or state actors for that matter] plan to use information technology, automated tools and identifiable targets may be observable and to some extent, predictable.'[27] This conclusion is borne out by history: Bletchley Park's wartime experiences and the JIC's Cold War performance demonstrate the importance of regularly monitoring and updating the progress of events, even if no predictive element is utilized; in other words, one of the JIC's key outputs was to describe the evolution and meaning of events, as much as it was to predict what might happen next.

What does this mean for the cyber field? At a strategic level, it signifies the importance of monitoring both the offensive capabilities of other nations and groups, but also in developing an objective assessment of defensive mechanisms.[28] One of the frequently cited reasons for being surprised is a misunderstood and incorrect view of how a country's defences stack up against an aggressor's power. The importance of strategic analyses is that they provide time for action to be taken; the difficulty, as one former CIA member has written, is that 'analysts must issue a strategic warning far enough in advance of the feared event for officials to have an opportunity to take preventive action, yet with credibility to motivate them to do so'.[29] The problem, certainly if the lessons of history are anything to go by, is that it is often very difficult to predict when a troop exercise on a border actually becomes the first move of an invasion. In the cyber context, for instance, is it possible to know whether attacks that try to expose vulnerabilities at different margins of a network are simply attempts to obtain access to information, probes to seek out vulnerabilities that could be exploited to create strategic effects, or smokescreens to divert attention from what is really planned?

The critical element of the tactical and operational levels is the ability and speed of response. Just as in the Cold War, simulations and exercises are valuable, as are detailed predictions as to

what might happen. This is particularly the case given that the decision to 'go' might be taken by one or two people (in which case advance intelligence warning is all but impossible) and that, once taken, operational effectiveness in the cyber age will be potentially instantaneous. A notion shared by conventional and cyber examples, is that 'if we accept the fact that warning is bound to be ambiguous then we must be capable of reacting repeatedly to false alarms without commitment to war'.[30] The intelligence producer-consumer relationship is critical here, not only for the manner in which warning is communicated to the policy realm, but also in the way in which it is received and responded to.[31] Betts has argued that the Cold War environment did not lend itself to this sort of fluidity: 'the only way to hedge is to have a military bureaucracy and strategic community capable of sensitivity, creativity, and quickly adaptive innovation. This is almost a contradiction in terms'.[32]

An ability to integrate both strategic and tactical intelligence is vital: thus, the government should be 'interested not only in explicit warnings but in general information about attacker intentions and capabilities'.[33] Does this mean that a JTAC-style outfit is the answer—an entity that includes seconded individuals from every relevant arm of the government working on cyber matters, designed to provide both strategic and tactical assessments? What it does suggest is that a strategy is extremely important, not only in discovering the attack and ensuring that all relevant bits of the state (including non-state responsibilities) work together, but also in terms of co-ordinating the response. This is something that former Director of National Intelligence (DNI) and Director of the National Security Agency (NSA), Mike McConnell, has argued the United States desperately needs.[34]

intelligence in the cyber era has to be both defensively and offensively aligned

The overarching lesson from examples of technological surprise is to avoid underestimating the opposing side's advances, particularly that novel ways might be sought to overcome obstacles. As the Germans discovered with their Enigma machine, just because you cannot defeat it, it does not follow that no one else will be able to either. The pace of technological change in the cyber domain has been frightening. Perhaps here, more than any other instance in history, an examination of internal vulnerabilities and defenses is critical. To ensure the best possible defense the lesson is that intelligence sharing is vital. In the cyber context this takes on a further dimension given the greater involvement of non-state actors to whom the state owes its allegiance. In this way not only is it important for the different strata of the government to work in unison, but it is equally as imperative to enable and provide a forum to share potentially classified information with those that are to be defended, be they banks, transportation services, critical infrastructure, etc.

The underlying theme, then, is that to be effective, intelligence in the cyber era has to be both defensively and offensively aligned. This last point is particularly valid in that the ability to respond is just as important in the cyber era as in conventional military examples. On the assumption that deterrence has failed and an attack has taken place, the manner and timeliness of the response will be critical. In the cyber field this might be a defensive reaction—stopping further penetration or attempting to discover the identity of the perpetrator—or it might be offensive, anything from a counter-cyber-attack to a military rejoinder. One lesson from conventional warfare is that it is particularly difficult to 'keep in constant readiness defenses strong enough to deter and, if need be, defeat a sudden attack'.[35] Presumably 'readiness' is less of a problem in the cyber era but is exacerbated by the proliferation of threats, possible attack routes and array of necessary responses.

THE CENTRALITY OF ATTACK MITIGATION: THE EXAMPLE OF THE UNITED KINGDOM

Unlike other forms of warfare, in the cyber domain the attacker has an overwhelming advantage: put simply, they can choose when, where and how to attack; the chances of being detected in advance are slim; and the speed with which an attack can take place suggests that they are unlikely to be interrupted. Furthermore, traditional notions of deterrence—in which the perceived response to an attack is enough to prevent it—are almost certainly inapplicable.[36] The central issue comes back to the original conception of a 'threat'. Williams, Shimmeall and Dunlevy argue that 'threats are inextricably linked

to vulnerabilities … in the realm of cyber-space, vulnerabilities are inescapable'.[37] Given the strategic advantage then, and the fact that pre-emptive intelligence of an attack is extremely difficult to obtain, it follows that the most important contribution is to be made once an assault is underway in the form of attack mitigation.

Klaus Knorr wrote that 'insurance takes three forms … the first is to provide for adequate military deterrence and defence; the second, to pursue foreign policies that do not unnecessarily incite the aggressiveness and despair of other states … the third form of insurance accepts the possibility of being taken by strategic surprise and is intended to reduce its impact'.[38] The British government's 2010 National Security Strategy installed cyber matters as one of four top priority concerns for the United Kingdom.[39] The Strategic Defence and Security Review (SDSR) subsequently provided £650 million of funding to tackle cyber issues, and the 2011 Cyber Security Strategy (CSS) focused on how this money should be spent.

At the heart of the CSS is an unstated but implicit recognition that cyber-attacks are impossible to avoid completely. Indeed, GCHQ has estimated that 80% of successful attacks in the United Kingdom could be prevented by 'simple computer and network "hygiene".' Thus, it is the remaining 20% to which GCHQ devotes its attention.[40] The implications are enshrined in one of the CSS's four objectives: 'making the UK more resilient to cyber attack and better able to protect our interests in cyberspace'. Although there is mention of the necessity to 'improve our detection and analysis of sophisticated cyber threats', the impression conveyed is that the cornerstone of the CSS is 'resilience' or mitigation—the ability to cope with, absorb, and continue in the aftermath of an attack (a variant of "graceful degradation").[41]

The means to achieve this are contained within the CSS's three other objectives, which center on the necessity to present a united front, not only within the governmental machinery, but also between the government and the private sector. In the short period since the CSS was formulated, there have been a plethora of new organizations created to foster collaboration, ranging from the tri-Service 'Joint Cyber Unit', hosted by GCHQ; the Centre for the Protection of National Infrastructure (CPNI), which is accountable to the Security Service; the National Cyber Crime Unit, soon to be based within the new National Crime Agency; through to the UK National Computer Emergency Response Team (CERT); the Cyber Incident Response scheme; and the Cyber-Security Information Sharing Partnership (CISP).[42] Most recently, in July 2013, the 'Defence Cyber Protection Partnership' was announced. A collaborative effort between the government and leading defense companies to improve 'collective defenses'.[43]

Crucial to the British response, and symptomatic of the need to pool information, are two bodies: (i) the Office of Cyber Security and Information Assurance (OCSIA), which resides in the Cabinet Office and liaises with other departments to provide strategic direction and coordination. (ii) The Cyber Security Operations Centre (CSOC), which is located in GCHQ but not technically a part of it (in a similar vein to JTAC and the Security Service). It draws its membership from secondments from other government departments and is intended to monitor cyberspace, coordinate incident responses, and understand the nature of any attack conducted.

Central to protecting the UK—or, by implication, any nation—is intelligence. The majority of the £650m allotted to cybersecurity has been distributed to the intelligence community, with GCHQ receiving the greatest share. In addition to providing tactical intelligence on evolving threats and means to cope, the community also provides strategic assessments of different states' cyber capabilities, particularly Russia and China. Yet, the role of the intelligence agencies is not restricted to protecting networks and thwarting plots. As the Intelligence and Security Committee wrote in mid-2012, 'while attacks in cyberspace represent a significant threat to the UK, and defending against them must be a priority, we believe that there are also significant opportunities for our intelligence and security Agencies and military which should be exploited in the interests of UK national security', including: active defense; exploitation; disruption; information operations; and military effects.[44]

For the United Kingdom then, the CSS provides a template from which structures have emerged, designed to foster collaboration and achieve common security. That is not to say it has

been without its critics, and one accusation has been the way in which it has lofty aspirations but little in the way of detail for how they should be achieved. Central to the strategy is the role of intelligence, and this is 'intelligence' in its widest meaning, encompassing everything from pre-emption, information assurance, and capacity building, to attribution and offensive action.

The overarching lesson from the history of surprise attacks is that surprise should not be unexpected. Admittedly contradictory, what this means is that omniscient intelligence is impossible and so surprise will always be conceivable. In the cyber era this conclusion is magnified: almost by definition every successful cyber attack is the result of surprise of one sort or another. The key, then, is in resilience and attack mitigation. Cyber attacks are a feature of modern life and, provided that all arms of the state (government and otherwise) are in communication, then surprise can be managed and expectations of the unexpected will become the norm. ❇

NOTES

DISCLAIMER: This paper is drawn only from released official records and published sources and the views expressed do not represent the views or carry the endorsement of the British Government. I am grateful to comments and feedback from Sir David Omand, Jim Wirtz, Thomas Rid, Dave Clemente and Dorothy Denning.

1 Dr. Michael S. Goodman is a Reader in the Department of War Studies, King's College London. He is currently on secondment to the Cabinet Office as the Official Historian of the Joint Intelligence Committee.

2 This paper is abstract in nature. For the most recent and most detailed account encompassing case-studies, see T. Rid, *Cyber War Will Not Take Place* (London: Hurst, 2013).

3 R. K. Betts, *Surprise Attack: Lessons for Defense Planning* (Washington, DC: The Brookings Institution, 1982) p.5 and p.6.

4 James J. Wirtz, "Theory of Surprise," in Richard K. Betts and Thomas G. Mahnken (eds.), *Paradoxes of Strategic Intelligence* (London: Frank Cass, 2003), pp.

5 These definitions draw upon Betts, *Surprise Attack*, pp.4-5. Also, C. Grabo, *Handbook of Warning Intelligence: Assessing the Threat to National Security* (Plymouth: Scarecrow Press, 2010) p.9; and J. W. Bodnar, *Warning Analysis for the Information Age: Rethinking the Intelligence Process* (Washington, DC: Joint Military Intelligence College, 2003) p.1.

6 R. K. Betts, 'Surprise Despite Warning: Why Sudden Attacks Succeed', *Political Science Quarterly* 95: 4 (1980-1), p.553. Emphasis in original.

7 Betts, *Surprise Attack*, p.15.

8 P. Morgan, 'The Opportunity for Strategic Surprise', In K. Knorr & P. Morgan (eds), *Strategic Military Surprise: Incentives and Opportunities* (London: Transaction Books, 1983) p.200. Emphasis in original.

9 Details are taken from *Avoiding Technology Surprise for Tomorrow's Warfighter: A Symposium Report* (Washington, DC: The National Academies Press). Available at: www.nap.edu/catalog/12735.html [accessed January 2013]

10 H.D.Kehm, 'Notes on Some Aspects of Intelligence Estimates.' *CIA Studies in Intelligence.* Available at www.cia.gov/csi/kent

11 These are from the perspective of the intelligence agency, from the target we might also add a 'declaratory intent'—that is, what is admitted (whether truthfully or otherwise) to.

12 For more see M. S. Goodman, 'Jones' Paradigm: The How, Why and Wherefore of Scientific Intelligence', *Intelligence and National Security* 24: 2 (2009), pp.236–56.

13 These are adaptations of the models described in R. M. Clark, 'Scientific and Technical Intelligence Analysis' In H.B.Westerfield (ed) *Inside the CIA's Private World: Declassified Articles from the Agency's Internal Journal, 1955–1992* (New Haven: Yale University Press, 1995) pp.294–7. The word 'capability' is deliberately vague, encompassing everything from a general to a specific capability.

14 Betts, 'Surprise Despite Warning', p.566.

15 R. Wohlstetter, *Pearl Harbor: Warning and Decision* (Stanford: Stanford University Press, 1962).

16 For instance, see the new preface to E. Kam, *Surprise Attack: The Victim's Perspective* (London: Harvard University Press, 2004). Also, C. F. Parker & E. K. Stern, 'Blindsided? September 11 and the Origins of Strategic Surprise', *Political Psychology* 23 (September 2002), pp.601–30.

17 In its simplest form: requirements and priorities; collection; analysis; dissemination.

18 D. Omand, *Securing the State* (London: Hurst, 2012).

19 K. Knorr, 'Lessons for Statecraft', In K. Knorr & P. Morgan (eds), *Strategic Military Surprise: Incentives and Opportunities* (London: Transaction Books, 1983) p.247.

20 Rid argues that 'no attack [has] damaged anything beyond data' thus far. T. Rid, 'The Obama Administration's Lousy Record on Cyber Security', *The New Republic* (6 February 2013).

21 'Allied Command Europe Counter-Surprise Military Alert System', 22 October 1956. Available at: http://www.nato.int/nato_static/assets/pdf/pdf_archives/20121128_19561022_NU_SHAPE-AG-1401-PP_ACE_Counter-Surprise_Miltiary_A.pdf [accessed February 2013]

22 Cited in M. W. David & K Sakurai, 'Combating Cyber Terrorism: Countering Cyber Terrorist Advantages of Surprise and Anonymity', *Proceedings of the 17th International Conference on Advanced Information Networking and Applications* (2003).

23 D. D. Clark & S. Landau, 'Untangling Attribution'. In *Proceedings of a Workshop on Deterring Cyberattacks: Informing Strategies and Developing Options for U.S. Policy* (Washington, DC: The National Academies Press, 2010) p.39. Rid makes a similar but more developed argument. *Cyber War Will Not Take Place.* p.140.

24 H. C. Deb 08 December 1941 vol 376 cc1355-64. Available at: http://hansard.millbanksystems.com/commons/1941/dec/08/prime-ministers-declaration#column_1358 [accessed February 2013].

25 For instance, see S. Goel, 'Cyberwarfare: Connecting the Dots in Cyber Intelligence' *Communications of the ACM* 54: 8 (August 2011).

26 Cited in P. Taylor, *Brits* (London: Bloomsbury, 2001) p.265.

27 David & Sakurai, 'Combating Cyber Terrorism'.

28 This latter point is particularly difficult, not least given problems of introducing regulation of the internet.

29 J. Davis, 'Improving CIA Analytical Performance: Strategic Warning', *Sherman Kent Center for Intelligence Analysis Occasional Papers: Volume 1, Number 1* (2002). See also A. L. George & J. E. Holl, 'The Warning-Response Problem and Missed Opportunities in Preventive Diplomacy', *A Report to the Carnegie Commission on Preventing Deadly Conflict* (1997). Available at www.carnegie.org [accessed February 2013]

30 J. Critchley, *Warning and Response: A Study of Surprise Attack in the 20th Century and an Analysis of its Lessons for the Future* (London: Leo Cooper, 1978) p.121.

31 S. Chan, 'The Intelligence of Stupidity: Understanding Failures in Strategic Warning', *The American Political Science Review* 73: 1 (March 1979), pp.171–80.

32 Betts, *Surprise Attack*. p.299.

33 Morgan, 'The Opportunity for a Strategic Surprise', p.209.

34 'Mike McConnell on how to win the cyber-war we're losing', *The Washington Post* (28 February 2010).

35 Critchley, *Warning and Response* p.5.

36 R. K. Knake, 'Untangling Attribution: Moving to Accountability in Cyberspace'. Statement prepared for the U.S. House of Representatives Subcommittee on Technology and Innovation, 15 July 2010. Available at: http://www.cfr.org/united-states/untangling-attribution-moving-accountability-cyberspace/p22630 [accessed February 2013].

37 P. Williams, T. Shimeall & C. Dunlevy, 'Intelligence Analysis for Internet Security', *Contemporary Security Policy* 23: 2 (2002), p.7.

38 Knorr, 'Lessons for Statecraft', p.264.

39 Cm 7953. *A Strong Britain in an Age of Uncertainty: The National Security Strategy* (October 2010). Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/61936/national-security-strategy.pdf [accessed February 2013].

40 Cm 8403. *Intelligence and Security Committee Annual Report 2011-2012* (July 2012). Available at: http://www.official-documents.gov.uk/document/cm84/8403/8403.pdf [accessed February 2013].

41 *The UK Cyber Security Strategy—Protecting and Promoting the UK in a Digital World* (November 2011). Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/60961/uk-cyber-security-strategy-final.pdf [accessed February 2013].

42 'Progress on the UK Cyber Security Strategy. Written Ministerial Statement'. (3 December 2012). Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/83756/WMS_Cyber_Strategy_3-Dec-12_3.pdf [accessed February 2013].

43 For more details see 'Keeping the UK Safe in Cyberspace'. Available at: https://www.gov.uk/government/policies/keeping-the-uk-safe-in-cyberspace [accessed February 2013].

44 'Ministry of Defence joins firms to tackle cyber threat', *BBC News.* Available at: http://www.bbc.co.uk/news/uk-23191891 [accessed July 2013].

45 *Intelligence and Security Committee Annual Report 2011–2012.* This includes descriptions for these offensive operations.

The Cyber Pearl Harbor Analogy: An Attacker's Perspective

Emily O. Goldman, John Surdu, & Michael Warner

Emerging technologies have had sweeping socioeconomic effects on states and societies. In recent decades, the comparatively low cost and availability of new information technologies have changed how people create, share, protect, and store data—and ultimately wealth. The effects often have been disruptive, initiating new lines of operation for businesses and exposing new weaknesses for competitors and adversaries to exploit. A "cyber" domain has emerged, an easily accessed, semi-ungoverned space that has become the newest front for military and economic confrontation.

WHY PEARL HARBOR?

On 11 October 2012, then-Secretary of Defense Leon Panetta publicly warned that the United States is vulnerable to a "Cyber Pearl Harbor." To many this conjures up grainy newsreel footage of burning battleships and the nation's entry into World War II. The Pearl Harbor analogy, from the U.S. perspective, evokes a devastating bolt from the blue that leaves an indelible imprint on the popular psyche. "Remember Pearl Harbor!" in this context is a call to mobilize support for increased cyber preparedness.

Our analysis of the Pearl Harbor analogy flips the perspective and proceeds from the viewpoint of the attacker. Some might argue that the warning klaxon of impending cyber attacks against the United States (not just on its military, but to its critical infrastructure) is alarmist and overblown. We are not so confident. Conditions exist on the world stage today that bear similarities to the East Asian crisis in the 1930s, which in turn led to the Japanese attack on Pearl Harbor in 1941. Such conditions could entice an adversary to strike a similar, disabling blow against the United States in the hope of a quick victory that presents America with an undesirable strategic *fait accompli*.

STRATEGIC CONTEXT FOR JAPAN'S ATTACK ON PEARL HARBOR

James Wirtz argues in his contribution to this anthology that Pearl Harbor was not in fact a "bolt from the blue," but rather a logical outcome of Imperial Japan's long-term strategy to solidify its hold on its Pacific empire. Japanese expansionism focused on establishing an exclusive sphere of influence—the Greater East Asia Co-Prosperity Sphere. By 1941, Japan's aggression in China and its larger aims in the Southwest Pacific were hampered by President Franklin Roosevelt's harsh sanctions, which Tokyo saw as tantamount to economic warfare. Japanese naval strategists thereafter sought to blunt Washington's ability to frustrate Tokyo's aims. Since America had moved aggressively to frustrate Japan's military aims in China and cripple its economy, the Japanese reasoned they had to hit back.

Japan's top naval strategist, Admiral Yamamoto, understood the risks of unleashing America's industrial might, which he had seen firsthand as a young officer posted to the United States. For Japan to win and retain the upper hand, he believed, it was necessary to strike an early, decisive blow—one that would preclude the possibility of the Americans going on the offensive while Japanese forces consolidated a defensive perimeter in the Western Pacific. Japan would have to eliminate U.S. forces in the Philippine Islands (astride Japan's desired supply routes) and crush the U.S. Pacific Fleet near Hawaii to prevent its advance toward Japanese home waters before Japan was ready for the great clash of battleships that, it was thought, would decide the struggle once and for all.

Yamamoto's strategic objective was not to conquer the United States or even to seize (much) American territory, but rather to delay the inevitable American counteroffensive. He judged that destroying the Pacific Fleet's offensive power, even temporarily, would allow Japanese forces to take control of oil supplies in the Dutch East Indies and erect a barrier chain of island bases, thereby enabling Japan to delay the Pacific Fleet in its westward course and perhaps even force negotiations from a position

of strength. A model for the attack was Germany's successful *blitzkrieg* strategy in France: hit hard and demoralize the adversary so its people would reject a long and costly war. It would take years for the United States to recover, Yamamoto hoped, and by then it would face a *fait accompli* with Japan's control extending from the Indian to the Pacific Oceans.

The Imperial Japanese Navy's Striking Force easily reduced the Pacific Fleet's old dreadnoughts on Battleship Row to smoking hulks, but missed the Fleet's aircraft carriers and left its fuel depots and drydocks almost undamaged. Yamamoto had taken a huge risk in mounting the Pearl Harbor operation—which could have been a tactical disaster for the Japanese fleet. American commanders across the Pacific expected war and already had their forces on alert. But local mistakes by commanders in Manila and Pearl Harbor ensured the Japanese had tactical surprise in both locations. The Japanese succeeded in destroying most U.S. fighter planes on the ground at Pearl Harbor, but the carrier-based air armada that struck Hawaiian targets that morning still took 10% casualties—a pretty high rate by World War II standards. Similar attacks by the Japanese in the Philippines that same day garnered similarly mixed results. Indeed, American bombers kept flying from the Philippines for weeks afterward attacking Japanese military targets.

adversaries who feel their backs are against the proverbial wall might lash out in new and unexpected ways

Yamamoto's aircraft carriers by then had run amok, as he predicted, supporting the conquests of the East Indies, Malaya, Singapore, Guam, and northern New Guinea, and raiding targets as far afield as Australia and India. Yet even then the timetable was slipping. In the Philippines, the Japanese quickly took control of the air and sea but had troops tied down for months trying to defeat the American and Filipino defenders. Even before these forces surrendered, American carrier-launched bombers had mounted a militarily trivial, but psychologically stunning, raid on Tokyo itself.

Pearl Harbor was a Pyrrhic victory for Tokyo, accomplishing some immediate goals but swiftly uniting the American people around calls for revenge. The surprise and scope of the defeat focused Washington's war aims on ending Japanese militarism, and ensured that the nation's determination would burn on for years. Yanked back to the grim imperative to sink the Pacific Fleet's aircraft carriers, Yamamoto and his staff began planning new operations to lure the Americans into a decisive battle, first in the Coral Sea and then near a remote atoll known to a handful of pilots and mariners as Midway Island. Together these battles certainly proved decisive, but not in the way Japanese planners wanted or expected. The Imperial Japanese Navy lost the cream of its pilots and two-thirds of its aircraft carriers within a month and forfeited the strategic initiative permanently. The United States could replace its naval losses and more. After Midway, Tokyo could only hope the Americans might run out of patience before they defeated Japan.

CYBER AS A TOOL OF CHOICE

Japan launched a war against the United States in 1941 because its leaders felt they were running out of time to respond decisively to President Roosevelt's sanctions. Japanese planners had powerful forces at their disposal but also knew their plans had to compensate for serious resource deficiencies. Is there a lesson here for our time? Perhaps it is that adversaries who feel their backs are against the proverbial wall might lash out in new and unexpected ways. We know what that meant in 1941, but what does it mean today?

Powers with ambitions and a high tolerance for risk—especially those that perceive the United States to be thwarting their ambitions—could seek ways to constrain our power to act. This includes impeding our ability to operate militarily, while also delivering a psychological blow to the American psyche. This logic applies in cyberspace as well. A cyber attack could have the virtue (from an adversary's perspective) of being more damaging to the United States' capacity to respond in the physical domain, and yet also be less "mobilizing" than, for example, a terrorist attack with mass casualties in an American city. A cyber attack might thus reduce America's ability to respond, while dampening (if not eliminating) political pressures to respond. This is the type of effect that a cyber-savvy adversary with outsized goals could find attractive.

Cyber activities can fit conveniently into an adversary's strategy to counter U.S. conventional military capabilities, making cyber weapons a tool of choice for potential U.S. adversaries. This is so for several reasons. First, cyber operations require skill and technical acumen indeed, but not the level of resources, manpower, and training required to leverage industrial-era technologies into sustained, conventional military forces. Second, the U.S. as an open society is easier to attack—and more vulnerable—in the cyber domain than it is in the land, maritime, and air domains. Third, cyber tools empower and enable even relatively unsophisticated actors to project power and operate deep within America's physical territory, to say nothing of its "virtual landscape." Fourth, and finally, cyber capabilities can be employed to achieve impacts like those of earlier "breakthrough" technologies that, in the past, undermined the existing advantages of even leading military systems. Potential adversaries can now avoid challenging us in areas of traditional military dominance and instead exploit our real and perceived vulnerabilities with cyber attacks launched against our critical infrastructure.

If an adversary's objective is to convince Washington to leave it alone, or to allow it to pursue its aims against its neighbors, then Admiral Yamamoto's intellectual heirs in such a situation could be tempted to mount a quick strike—against our mobilization and logistical network—that would keep the United States at bay. Potential adversaries have invested in anti-access and area-denial capabilities, as well as in other asymmetric means to counter traditional U.S. strengths and to prevent America from projecting power abroad. They could prepare the future cyber battlefield now by stealing intellectual property, conducting industrial espionage, and exploiting government networks and those of our defense, financial, and communication industries. Through intelligence, surveillance, and reconnaissance against U.S. and allied networks, they could gain penetration and establish persistent access. We are witnessing such activities today, as can be verified from perusing the continuous and alarming public statements uttered by a host of independent computer and software security experts.

ADVERSARY PERCEPTIONS AND CALCULATIONS

What would such an attack actually gain, and what might make it seem worth the attendant risks? To recall our Pearl Harbor example, Japanese military planners had no illusion they could fight and win a war of conquest against the United States. They preferred not to fight the Americans at all, but felt the proverbial dice had already been tossed by President Roosevelt's economic sanctions. With conflict seemingly inevitable, Tokyo sought to keep Washington's only offensive force—the U.S. Pacific Fleet—as far away as possible from the regions that Japan desired to control. In short, in an aggressor's calculus, if the United States might be induced to stay out of a regional conflict, then the aggressor keeps his gains. Once he has gained what he wanted, an aggressor might even have the ironic temerity to call on the international community to intervene to stop U.S. "pressure and intervention."

A *fait accompli* is a half-way step toward a larger war. It promises a greater chance of political victory than quiet diplomacy, but it also raises the risk of violence. The acting side moves without warning, presenting the other side with an accomplished fact. Retreat means losing face, while standing firm most assuredly means collision.[1]

Do people actually think this way? Saddam Hussein of Iraq certainly did in 1990. He mounted a surprise mobilization of his best divisions and then invaded neighboring Kuwait as soon as his forces were ready. Kuwait fell to Iraqi troops in mere hours, giving Saddam an oil-rich "19th province" with a fine harbor on the Persian Gulf—and changing at a stroke Iraq's strategic

position with respect to Iran, its enemy in a grim eight-year war that had just ended. But Saddam's gambit failed, and his forces were driven from Kuwait.

Can a new aggressor hope to achieve surprise in our modern age? Would not its moves be telegraphed in advance during a period of rising tensions? The likelihood is not high—it is virtually certain. Every surprise attack over the last century had to overcome the peril of discovery. Indeed, there is no example of a significant surprise attack in living memory in which the victim had no warning of what could happen. Surprise works not because the victim has no inkling of peril but because he believes the peril is not *imminent*. It is not comforting to reflect that surprise probably plays a larger role in cyberspace than in any other domain. The digital battlefield provides unprecedented opportunities for surveillance, deception and surprise. Devastating attacks on critical infrastructure can be launched with little or no tactical warning (although presumably few actors can do this yet).

low-grade cyber conflict has been increasing for years

We believe low-grade cyber conflict has been increasing for years, that cyber interference with U.S. interests (including our economy and future economic competitiveness) are occurring daily, and that we are late-comers to the cyber warfare game. Our opinions, however, are less important than others' perceptions of the United States. Viewing Pearl Harbor from the attacker's perspective compels U.S. policymakers who are grappling with the tactical, operational, and strategic implications of conflict to include a cyber dimension to their calculations, and to reflect on how our actions could influence adversary considerations and planning.

Take, for example, the creation of U.S. Cyber Command in 2010. This was meant to address a national, strategic vulnerability to cyber attacks. In the 1930s, the U.S. feared the rise of Japanese militarism and quietly began preparing for war. But could the establishment of U.S. Cyber Command have a similarly unintended consequence? Could its creation pose a threat to our adversaries? Some have worried that the creation of U.S. Cyber Command could have the undesired, long-term side effect of furthering a world-wide rise in cyber militarism. An adversary might believe that U.S. Cyber Command is a primarily offensive organization training weapons on them. Such an adversary might see the need to strike against the U.S. economy sooner, rather than later, before the Cyber Command is able to strengthen America's defenses against cyber attacks. We should consider how the creation of U.S. Cyber Command could be shaping the offensive calculations of our adversaries, and whether, like Japan in the 1930s, they might consider a "bolt from the blue" in order to temporarily blunt U.S. capabilities.

WARNING AND RESPONSE

The Obama administration signaled its attitude toward cyber attacks when the White House released its vision for the future of cyberspace. "When warranted, the U.S. will respond to hostile acts in cyberspace as we would to any other threat to our country," it said, adding that such responses included "all necessary means," including military ones. Complicating the matter of responding to cyber attacks, however, is the debate over what constitutes an act of war in cyberspace. At the current time, observers inside and outside the government disagree about what the Department of Defense (DoD) can and should do in response to malicious activity in the cyber domain. There is no consensus on what acts, including threats against the United States or responses to those threats, would constitute a hostile act. The same disagreement manifests itself at the international level, where debates continue over what the threshold is for declaring a cyber attack an act of war.

With these geostrategic and policy considerations as a backdrop, what lessons can the U.S. draw if we decide that the Pearl Harbor analogy is a useful one to ponder from the attacker's perspective? What should we build in order to apply these lessons?

Defensible Architecture

Defenses make a difference; even bad ones can slow an attack's momentum, and strong defenses can prevent defeat. The U.S. Navy said as much over a year before Pearl Harbor. Navy brass had opposed President Roosevelt's decision to move the Fleet from San Diego to Hawaii in 1940 (Roosevelt wanted to send

Japan a warning), so one lesson is to reduce one's defensive footprint (or the "attack surface") for an adversary to target.

We must assume our networks have been penetrated already. We need layered defenses to secure them and prevent strategic surprise. This begins with overall network hygiene. It requires sensors within and at the boundaries of our networks to act as scouts, screening for adversary attack methods and detecting and eliminating adversary activity within our networks. Scouts can also conduct counter-reconnaissance to deny the adversary the ability to map our command and control nodes or to gather information about our defenses. We must be able to see in order to defend. Yet with 15,000 network enclaves in DoD alone, this is extremely difficult. DoD is working to reduce its own attack surface by limiting the number of its connections to the Internet backbone. Even if we could see all our networks, however, most cyber operators today are busy applying patches and configuring devices. We must automate updates to our systems to free up humans to hunt for adversary malware inside DoD networks. Still, the nation, and not just the Department of Defense, is under incessant cyber attack. As a nation, we must reduce our vulnerability to them. DoD can assist through information sharing; provision of assistance teams to operate under the control of the supported organization; and deployment of defensive capabilities in or at the boundaries of others' networks.

Situational Awareness

The strong defenses at Pearl Harbor lacked situational awareness and tactical control. Early detection of Japanese aerial and submarine scouting of the harbor on 7 December should have prompted the base to go to battle stations. As it was, however, the Army and Navy failed to coordinate their scouting and watches, and clear indicators went unheeded. A common operational picture might have prevented this. Had the radar system sent to guard Pearl Harbor been fully operational, for example, the Japanese attack could have been blunted.[2]

In cyberspace as well as air operations, a common operational picture (COP) is crucial. Such a COP is difficult to gain in cyberspace, because there are so many dynamic data points. Still, the data points likely exist somewhere, and political, cultural, and legal considerations could allow for an early warning system for cyber attacks. Cyber intelligence must involve both cyber and non-cyber sources in order to build a comprehensive operational picture. As with nuclear weapons, this early warning system must involve close integration with our allies.

Active Defense

What would have been the U.S. response if the right dots had been connected in early December 1941? International law permits nations to conduct preemptive strikes. Seeing six Japanese aircraft carriers north of Hawaii and headed for Oahu at top-speed on December 6, for example, U.S. Army bombers might well have been ordered, with clear legal justification, to launch attacks against that fleet. There are no clear norms, however, for preemptive strikes to blunt or prevent cyber attacks against our national infrastructure. Nonetheless, after thwarting an adversary attack, we must have the capability and authority to maneuver within our own networks, in neutral networks and, if directed, into adversary networks, to conduct operations that neutralize and disrupt the adversary's ability to conduct follow-on cyber operations.

Intelligence

Japanese operational deception worked before Pearl Harbor. U.S. Army analysts had ample diplomatic signals intelligence from reading Japan's foreign ministry ciphers. But U.S. Navy analysts misread the SIGINT clues they possessed (and failed to realize how many indicators they lacked), partly because the Japanese fleet practiced simple but effective deception and denial methods. Better management of intelligence analysts in Washington and Hawaii might well have revealed additional clues to Tokyo's intentions, and spotted the Japanese deception efforts—providing another vital indicator of impending hostilities.

Intelligence is a process akin to assembling a mosaic. Then and now the process must be comprehensive, active, and well-managed—and protective of the civil liberties and privacy of the American people. Today we operate in a telecommunications environment that is perhaps the most complex artifact ever devised. We cannot stop the advance of technology, but we must make every effort to use technology to improve our intelligence without compromising our values.

Resilience

The purpose of Japan's Pearl Harbor attack was to delay an American response, not to achieve ultimate victory. How long might it take us to recover from a massive attack on our critical infrastructure today? Would that buy an adversary time to operate without fear of U.S. retaliation? Is there a contemporary analogue to a decisive blow against Battleship Row—or are we more secure by having a distributed infrastructure? Building resilience and redundancy into our critical infrastructures is essential. In cyberspace our "Battleship Row" might consist of numerous national systems and institutions critical to our economy, such as, but not limited to: undersea cables, Wall Street, power grids, water supplies, classified networks, electronic voting systems, banking systems and electronic funds transfer (those that enable Internet commerce, for instance), and a very heavy reliance on a single operating system (i.e., Windows). We have seen recent examples of large consequences brought about by inadvertent interruptions in these systems. We can only imagine what would be the effects of intentional disruptions.

Attribution

Attribution was not an issue at Pearl Harbor and this represents a key difference between cyber attacks and traditional attacks. Attribution is commonly cited as a reason that cyber warfare favors the attacker; it can be difficult to determine the attacker unambiguously due to deception, pre-deployment of attack tools, and attack through convoluted paths. It is also easier in cyber warfare to hide the buildup of "cyber ships, tanks, and aircraft." Cyber tools can be developed within isolated networks, denying us intelligence about enemy capabilities. These characteristics may make cyber attacks a good risk for an attacker to take precisely because they create complications for response.

debates continue over what the threshold is for declaring a cyber attack an act of war

Cyber and Combined Arms

On the other hand, cyber attacks could invite a stronger, kinetic response that the attacker may well regret. Anyone who chooses to employ cyber attacks against the United States must take into account the possibility of a response that might not necessarily be executed in cyberspace. Today, a kinetic response to a damaging or crippling cyber attack might be seen as overreaction; however, as norms of behavior at the nation-state level in cyberspace are formed, such responses may become more acceptable.

At the national level, we need to synchronize cyber operations with traditional military activities and other initiatives. National leaders need to understand the comparative advantages and disadvantages of achieving an effect by cyber means versus conventional means, the risks of escalation, and how to generate mass and coordinate movements across the cyber and physical domains. We understand conventional combined arms operations like AirLand Battle and AirSea Battle—now we may need to develop "Cyber-Kinetic-Diplomatic-Law Enforcement Battle."

Trained and Ready Forces

We had little up-to-date experience in maneuvering large sea, air, and ground forces at the outset of World War II. The forces at Pearl Harbor had never been bombed by modern aircraft, and did not know how to prepare for air raids. Proper dispersal of aircraft and warning (possible even without radar) would have reduced the accuracy of the raids and downed more attackers. Our infrastructure remains woefully unprepared for cyber attacks. Yet we have an advantage in that the incessant activities against our networks are helping us prepare. The constant cyber skirmishing that characterizes the daily struggle on our national infrastructure helps build experienced cyber warriors and more defensible systems.

CONCLUSION

Cyber warfare has already begun, at least at the level of spying, skirmishes, and patrols. One can argue our options, but we would be unwise to ignore the possibility of sudden major cyber attacks. We currently fall short in critical respects that could have repercussions for both the defense of the nation and the ability of Combatant Commanders to execute their operational plans. Our nation's capabilities in cyberspace are strong but limited. Small comfort lies in the realization that other countries probably lag behind us in adapting military operations for the cyber domain—as they were in the air and sea domains a century ago. We are racing potential adversaries

to learn how to fight in, through, and alongside cyberspace, and new contenders will continue to enter the race.

The ease of applying mass and achieving surprise in the new cyber domain means that we now have some peer or near-peer competitors in this space. Threats to U.S. national and economic security in cyberspace are increasing in complexity and destructiveness, and given the lack of traditional warning and the lack of immediately visible consequences of cyber exploitation and disruption, resources must be directed to address critical shortfalls to protect the nation. We must maintain our momentum, improve our defenses, and close the gaps outlined above that threaten our well being and strategic security. Actions are needed to manage the current cyber risk and allow for long-term development and sustainment of cyber defenses. Global circumstances require agile and technologically advanced cyber capabilities empowered to respond to threats to the United States beyond the borders of U.S. military information systems. We must gain greater global cyber situational awareness, exercise authorities to defend the nation, adequately command our cyber workforce, build a defensive cyber architecture, and train cyber teams to defend against growing cyber threats.

The Japanese gamble almost worked at Pearl Harbor. The surprise attack could have disabled the U.S. Pacific Fleet for a year or more, giving Japanese forces time to dig in for a lengthy conflict. But Japan's knockout blow didn't knock us out, and it ensured the United States would fight to the end of Japanese militarism. Pyrrhic victories can be very expensive to those who win them but ultimately lose the conflict. They can cost just as much to those defeated at the outset who nonetheless ultimately triumph in the end. For these and other reasons such conflicts are best not fought at all. That means, however, they must be deterred—even if doing so costs considerable resources and seems to risk "militarizing" those assets and interests that one feels beholden to defend. ❇

NOTES

1 Alexander L. George, "Strategies for Crisis Management," in Alexander L. George, *Avoiding War: Problems of Crisis Management* (Boulder: Westview, 1991), 377–94, 382–83.

2 Gordon W. Prange, *At Dawn We Slept: The Untold Story of Pearl Harbor* (New York: McGraw-Hill, 1981), pp. 495–502.

"When the Urgency of Time and Circumstances Clearly Does Not Permit...": Predelegation in Nuclear and Cyber Scenarios

Peter Feaver
Professor of Political Science, Duke University

Kenneth Geers
Senior Global Threat Analyst, FireEye

ABSTRACT

The speed with which a devastating cyber attack could strike the U.S. means there may be insufficient time for the traditional political-military decision-making process to work. During the Cold War, the U.S. faced a similar challenge in the nuclear domain. To meet this challenge, the nuclear command and control system granted a small number of senior military commanders a "predelegation" of authority to use nuclear weapons in the event of a sudden, catastrophic national security crisis. We review the lessons learned from nuclear predelegation and apply them to the cyber domain. We conclude that the cyber command and control system may need to consider similar predelegation measures in order to defend against some forms of hostile cyber attack. However, there are inherent risks in this policy that are also analogous to the ones confronted in the nuclear era. Counterintuitively, defending the nation against cyber attacks may demand patience rather than predelegation.

1. Introduction

In a formerly TOP SECRET document, entitled "Instructions for the Expenditure of Nuclear Weapons in Accordance with the Presidential Authorization Dated May 22, 1957," the U.S. military was notified that "When the urgency of time and circumstances clearly does not permit a specific decision by the President, or other person empowered to act in his stead, the Armed Forces of the United States are authorized by the President to expend nuclear weapons in the following circumstances in conformity with these instructions…".[1]

The importance and significance of this directive was underlined by the fact that President Eisenhower personally informed then-Secretary of Defense Thomas Gates that the President himself had written parts of it. Furthermore, Eisenhower informed Gates that "I cannot overemphasize the need for the utmost discretion and understanding in exercising the authority set forth in these documents. Accordingly, I would like you to find some way to brief the various Authorizing Commanders on this subject to ensure that all are of one mind as to the letter and the spirit of these instructions." Eisenhower's memo shows the U.S. national command authority wrestling with the thorniest of national security concerns: how to preserve political control when evolving technology and threats are pushing for faster and faster response. Today the national command authority is facing similar issues and perhaps policymakers and cyber commanders can learn from the efforts of earlier generations to adapt to the nuclear age. Although cyber conflict does not constitute the same kind of civilization-ending threat that global thermonuclear war poses, it may well demand changes to the way U.S. leaders manage national security affairs that will rival the changes wrought by the advent of nuclear weapons seventy years ago. In particular, just as nuclear weapons imposed unusually dramatic constraints on traditional command and control arrangements, it is likely that cyber conflict will strain existing command and control systems in new ways.

In this paper, we examine one specific parallel: predelegation policy, which granted lower-level commanders the authority to use special weapons under carefully prescribed conditions. Three features of nuclear war drove policymakers to consider and, in some cases, to adopt innovative approaches

to predelegation: 1) the speed with which a nuclear attack could occur, 2) the level of surprise that could be achieved, and 3) the specialized nature of the technology (which meant that only certain cadres could receive sufficient training to be battle-competent).

For each of these features, there is an obvious cyber analogue. In both nuclear and cyber defense, 1) defenders are under a great deal of pressure to act quickly, 2) they may be faced with conflict scenarios no one could have imagined, and 3) the nature of the fight requires a high level of training and technical expertise. As a result, and in both the nuclear and cyberwar cases, defenders may require some level of predelegated authority in order to act quickly and capably in defense of the nation.

Thus, the "letter and spirit" of Eisenhower's memorandum is also the topic of this paper: namely, the possibility that certain threat scenarios may oblige the National Command Authority to do something it would much prefer not to do: to authorize military action in advance, without knowing exactly when and how it will be used.

defenders may require some level of predelegated authority in order to act quickly and capably

2. Nuclear Predelegation

Early in the nuclear age, policymakers recognized a trilemma inherent in the nuclear revolution (Feaver, 1992).[2] The first two horns were the "always-never dilemma"—political authorities demanded that nuclear weapons *always* be available for use even under the most extreme conditions (even after suffering a surprise attack), while at the same time demanding that they would *never* be used accidentally or without proper authorization. Many measures designed to assure the "always" side of the dilemma posed risks for the "never" side, and vice-versa. The third horn of the trilemma was the desire that nuclear weapons enjoy the highest level of civilian control, far in excess of what was required for conventional military weapons and operations. Here, some measures designed to ensure strict civilian control tended to exacerbate the always-never dilemma. What happened in practice? In fact, the evolution of the U.S. nuclear command and control system was an ongoing set of compromises which balanced myriad risks against these three desiderata.

As the Soviet nuclear arsenal grew in size and lethality, the challenges of this trilemma became more acute. What if a sudden illness, a natural disaster, or a surprise military attack killed or incapacitated the President, and perhaps other senior leadership figures, before they could even begin to manage a war? What if tactical commanders received warning of an attack—or actually came under attack—and political authorities delayed in responding? Relative to certain weapons, this could create a "use them or lose them" scenario. What did we want U.S. nuclear commanders to do in these and other dire scenarios, and how could we ensure that they would not violate the principles of always, never, and civilian control?

One controversial measure designed to address these concerns was *predelegation of use authority* (hereafter, predelegation), in which the President spelled out carefully delineated procedures in advance that would authorize when and how nuclear weapons could be used by tactical commanders. Of course, some form of predelegation is as old as warfare itself. As Martin Van Creveld observed, even Stone Age chieftains wrestled with the challenges of command in war and part of their solution likely involved explaining to the other warriors what they should do under certain anticipatable circumstances (Van Creveld, 1985, pp. 5–9). For centuries, and before technological advances solved the problem of communicating at great distances, ground and especially naval commanders departed on their missions with orders that spelled out in greater or lesser detail what political authorities expected the commanders to do while out of communication range. Indeed, some form of predelegation is inherent in the President's function as chief executive officer; unless the President can delegate certain of his powers and duties, little in the country would ever get done.[3] Faced with the trilemma of always, never, and civilian control, U.S. national command authorities updated this familiar tool to the unfamiliar constraints of the nuclear age. It has long been known that, between the Eisenhower and Ford administrations, up to seven unified and specified commanders, at the three- and four-star levels, possessed the authority to launch nuclear weapons (Bracken,1983, pp. 198–199; Sagan, 1989, p. 142; Blair, 1985, p. 113). In 1950, CINC Strategic Air Command (CINCSAC) General Curtis

LeMay argued that senior officers must be able to act in the event Washington were destroyed by a surprise Soviet attack, and later believed that he had gained "de facto" authority (Feaver, 1992). In 1957, LeMay informed a presidential commission: "If I see that the Russians are amassing their planes for an attack, I'm going to knock the shit out of them before they take off the ground" (Kaplan, 1983, p. 134). His successor, CINCSAC General Thomas Power, informed Congress that he possessed "conditional authority" to use nuclear weapons. During the 1962 Cuban missile crisis, Supreme Allied Commander Europe (SACEUR) General Lauris Norstad was given prior authority to use nuclear weapons if Russia attacked Western Europe (Feaver, 1992).

The nature and scope of nuclear predelegation has been some of the most highly classified information in the U.S. nuclear establishment, so the public record is murky and filled with holes. Nevertheless, in the last 15 years, a number of documents have been declassified, filling in some of the gaps.[4] The most dramatic revelation was the declassification of information on "Project Furtherance," a plan that provided for "a full nuclear response against both the Soviet Union and China…." under certain circumstances, particularly "in the event the President has been killed or cannot be found."[5] In the memo dated October 14, 1968, President Johnson's advisors recommended changes to the existing authorities, specifically allowing for the response to be tailored either to the Soviet Union or China, limiting the response to a conventional attack at the non-nuclear level, and outlining these instructions in two rather than one document. The most recent revelations indicate that predelegation extended well beyond the use of nuclear weapons in a defensive role.

In 1976, the U.S. reportedly planned to revoke some if not all of the provisions for nuclear predelegation that it had established in the 1950s (Feaver, 1992). Currently, it is not publicly known whether any predelegation of authority to launch nuclear weapons continues to exist and, if so, under what constraints. However, based on recently declassified documents, it was clear into the 1980s that the threat of decapitation and the difficulty of maintaining connectivity with national command authorities during a nuclear war was still very much a problem that U.S. war planners were addressing—and that predelegation was at least one of the options under debate.[6]

2. 1 The Pros: Why Nuclear Predelegation

The primary benefit of predelegation is that it is a reliable workaround to the threat that the enemy might interdict communications between national command authorities and nuclear operators, which could decapitate the nuclear arsenal and render it impotent. Moreover, predelegation fixes this problem while simultaneously reinforcing the legal chain of command. The predelegated instructions take the place of the orders that the national command authority presumably would have given in the scenario if able to, thus making the actions legal.

Predelegation is preferable to presidential succession, which transfers all presidential authority to subordinate officials. The Constitution and the Presidential Succession Act of 1947 prescribe a cumbersome process of succession from the President, to the Vice-President, to the Speaker of the House, to the President *pro tempore* of the Senate, to the Cabinet officers in order of when each Department was established. But a nuclear war could kill many if not all of these civilians suddenly, or at least render them *incommunicado*. And given the secrecy and complexity of nuclear war planning, it is doubtful that more than a handful of these officials would be ready to manage a war, especially a nuclear war. In short, national security planners have good reason to fear that the constitutional line of succession would move too slowly during an extreme national security crisis.

A crisis-oriented alternative to succession is the "devolution" of military command, in which the President as CINC is replaced by the Secretary of Defense, immediately followed by the next highest-ranking military officer, and so on. However, it is highly likely that any practicable system of *de facto* devolution of command would quickly diverge from the *de jure* line of succession. Furthermore, devolution as a national plan may rest on shaky political and legal ground. It is doubtful that U.S. civilian leadership would ever agree to cede so much power to the U.S. military automatically, and the Supreme Court might not uphold it as constitutional. Finally, devolution of command creates the problem of "multiple presidents" if communication links with one or more of the individuals in the chain of command are reconstituted and then lost again as the crisis evolves.

Predelegation is on much stronger legal ground and is thus preferable to devolution of command. Predelegation gives conditional, *de facto* authority to certain trusted commanders while keeping *de jure* authority with elected civilian leadership. Moreover, predelegation allows for fine-tuned civilian control, since the predelegated authority can be as restrictive or permissive as desired. Thus, predelegation appears to reinforce civilian control of nuclear weapons. Last but not least, predelegation allows the president to reassert command and control if communications are restored.

predelegation poses a strain on civil-military relations

It is not enough to have policies and doctrine aimed at addressing the trilemma. The political authorities must also understand the doctrine and support those policies on an ongoing basis. Military doctrine without political buy-in cannot be sustained indefinitely. Over time, gaps are likely to emerge between what political leaders think is military doctrine and what military officers understand to be their doctrine. During a crisis, this lack of mutual understanding could lead to response failures or other breakdowns in command and control, proving disastrous for the nation.

In sum, compared with the alternatives and provided political authorities fully comprehend what they are doing, predelegation is simple and easy to implement. Building hardened command, control, and communication (C3) networks that would withstand every worst-case scenario imaginable would be prohibitively expensive, even if it were technologically feasible in the first place. Predelegation offers a ready stop-gap for unforeseen circumstances that could defeat U.S. C3 networks and is, by comparison, essentially free.

2.2 The Cons: Why Not Nuclear Predelegation

The predelegation of nuclear authority has an age-old Achilles heel: human nature. In order for the system to work in the extreme scenarios when it would be needed, it couldn't be stymied by technical measures that physically block use (such as Permissive Action Links or other coded systems that separate possession from ability to use).[7] Predelegation was intended as the solution for when all communication with political authority was broken. Therefore, a military commander possessing predelegation authority must also possess everything that he or she would require to give a legitimate launch order. Logically, a commander with predelegated authority must be able to make an unauthorized use look authorized to anyone downstream in the chain of command. Thus, predelegation favors the "always" side of the trilemma at the expense of the "never" and the "civilian control" sides. These risks are tolerable provided that the commanders honor the terms of their predelegated authority—that is, if they can be counted on to operate with complete integrity. Of course, the nuclear establishment invests extensive resources to ensure such integrity, but this risk is not inconsequential.

Predelegation seems to imply that *de jure* political control would give way very quickly to *de facto* military control, and that there would be some level of automaticity to nuclear retaliation akin to the interlocking mobilizations of World War I or to the Soviet Union's "Dead Hand" system (Lieber, 2007; Hoffman, 2010).[8] In short, predelegation poses a strain on civil-military relations. As personified by General Curtis LeMay and parodied in *Dr. Strangelove*, in war as in peacetime, civilian and military leaders may have different tendencies. Military officers, to protect assets, forces, or territory, may desire to employ nuclear weapons in preemptive or retaliatory action, even if the bombs explode over domestic or allied territory. They may feel a certain pressure to "use them or lose them." On the other hand, civilians might prefer to absorb

tactical military losses in the hope of achieving strategic gains, such as preventing an escalation of the conflict.

As a concept, predelegation is simple, but in practice it must be a highly complex mechanism. For example, how far down the chain of command should nuclear authority go? How wide should the latitude be, and how specific the instructions? It is hard to anticipate in advance what would be the preferred course of action under scenarios that can be only dimly imagined. In practice, for predelegation to be effective, prescribed conditionality would have to be balanced with implied flexibility. Yet it is undesirable to have too much interpretive latitude with nuclear weapons. And how public should predelegation policy be? Revealing some information helps deterrence, but revealing too much information gives the enemy opportunities to figure out how to defeat the system (Rosenbaum, 2009).[9] It is worth noting that presidential delegations of authority should be published in the Federal Register, but this never happened with nuclear authorities (Feaver, 1992). Finally, how should nuclear authority revert to civilian control? In theory, this should happen as soon as reliable communication with the President or his successor is restored, but in practice this would be difficult to accomplish during a rapidly unfolding crisis.

predelegation could apply both to offensive and defensive weapons.

In theory, predelegation could apply both to offensive and defensive weapons. However, the case for nuclear predelegation is much stronger for defensive weapons, such as air defense missiles tipped with nuclear warheads. Defensive weapons have a very short operational window to be effective, and the consequences of an unauthorized defensive use may be less severe than for unauthorized offensive use. Defensive nuclear weapons would explode primarily over U.S. and Canadian airspace. By contrast, offensive weapons would detonate on enemy territory, greatly increasing the pressure to escalate the crisis.

However, the fact that even *defensive* predelegation scenarios involved the territory of other states proved to be one of the most sensitive and difficult aspects of the policy. The declassified record shows that President Eisenhower reluctantly acquiesced to the predelegation policy, but he was personally most invested in dealing with the political challenge of pre-authorizing nuclear activity that would so directly affect our closest allies. In the notes from a formerly TOP SECRET meeting on June 27, 1958, "The President stressed the weakness of coalitions as bearing on this matter [referring to the predelegation of authority to fire nuclear air defense weapons]. He recalled that this was largely the secret of Napoleon's success, which was not seen until Clausewitz wrote about it. He recalled that Clausewitz had stressed that war is a political act—we must expect the civil authorities to seek control."[10]

3. Cyber Predelegation

The nuclear revolution began with an historic explosion in the New Mexico desert on July 16, 1945. By contrast, the cyber revolution slowly sneaked up on us. And while the Internet era has truly benefited the entire world, a looming downside is that the world may have grown too dependent on a technology that is highly vulnerable to attack. We are still at the beginning of the Internet era, but almost every kind of network-connected critical infrastructure has been threatened by hackers: air traffic control (Gorman, 2009), financial sector (Wagner, 2010),[11] elections (Orr, 2007),[12] water (Preimesberger, 2006),[13] and electricity (Meserve, 2007).[14] Over time, this problem may only get worse, as formerly closed, custom IT systems are being replaced with less expensive commercial technologies that are both easier to use and easier to hack (Preimesberger, 2006). National security thinkers rightly worry that militaries, intelligence agencies, terrorists, insiders, and even lone hackers may target such systems in the future.

Cyber weapons do not pose an immediate, apocalyptic threat on the scale of nuclear weapons; and for the foreseeable future, the always-never dilemma will not apply in the cyber domain exactly as it applied in the nuclear domain. Indeed, in the nuclear era, apart from the bombs dropped on Hiroshima and Nagasaki, the U.S. military only prepared for nuclear war—but never fought one. By contrast, the U.S. national security establishment (and even the private sector) is almost always under some form of cyber attack, even though many key players have scarcely begun to prepare for it. The U.S. may have a low tolerance for the kind of catastrophic cyberattack envisioned in worst-case scenarios, but we manifestly have a high tolerance for the low-level cyberattacks that occur every day.

Still, as the infamous Morris worm of 1988 and the more recent Stuxnet computer worm illustrate, there are reasons to worry about the intended and unintended effects of authorized or unauthorized use of cyber weapons (Broad & Sanger, 2011). Moreover, cyber has a novel dimension that dramatically intensifies the degree of concern about political control: the line between military/national security and civilian/commercial domains is quite blurry and activities in one domain usually seep over into the other, raising sensitive privacy and civil liberty concerns. On the notional spectrum from bayonets to ballistic missiles, cyber weapons are often considered to be closer to the ballistic missile end, requiring extraordinary command and control arrangements—not unlike nuclear weapons. However, all of these assessments are tentative, and still open to debate.

There are important analogs between nuclear attacks and cyber attacks: 1) malicious code travels across computer networks at lightning speed; 2) successful cyber attacks are often based on novel ideas; and 3) computer security is a complex, highly technical discipline. These three characteristics—speed, surprise, and specialization—may force national civilian leadership to give tactical military commanders a predelegated authority to operate in cyberspace, so that they are able to defend U.S. computer networks competently and successfully.

Yet the cyber challenge differs from the nuclear in two key ways—attribution and impact. Together, these point to the need for caution in adopting the nuclear era "fix" of predelegation. In cyberspace, it is often difficult to know with certainty who is attacking you, at least until a full-scope investigation is complete. This poses a significant obstacle to quick retaliation. There are analogous concerns in the area of nuclear terrorism, but for most of the Cold War, the attribution concern from state-based attacks was a secondary consideration. Likewise, if cyber attacks do not pose an existential threat to American society, they also do not pose the always-never dilemma. Therefore, it is politically fraught to assume the risks inherent in predelegation, and the benefits and requirements are more open to debate. Predelegation was quite controversial during the nuclear era, when the command and control exigencies made it seem even more necessary. Therefore, cyber commanders could have more difficulty than their nuclear predecessors in convincing political leaders of the wisdom of the predelegation option.

3.1 The Pros: Why Cyber Predelegation

First, it may take months or even years to plan a cyber attack, but once an attacker pulls the trigger, electrons move far more quickly than ballistic missiles—at close to the speed of light. In fact, even layered cyber attacks may unfold at such a high rate that predelegation alone is insufficient. For nuclear war, predelegation was deemed necessary to eliminate cumbersome interactions between national command authorities and tactical commanders. However, under most scenarios, tactical commanders would likely have enough warning to make their own deliberative response. With cyberattacks, the damage is often done before tactical commanders have a chance to collect evidence, evaluate data, and prepare a response. The cyber analog therefore might not be *predelegated authority to respond* but *automated authority to respond*. One of the primary fears of nuclear predelegation was that there would be an automatic response. But with cyber attacks, the minuscule time windows involved could make some level of automation inevitable.

Second, nuclear predelegation hedged against surprise attacks and unforeseen scenarios. Cyber attacks are also characterized by a high level of surprise. Information technology and cyber attacks are evolving at a blinding rate. It is impossible to be familiar with every hacker tool and technique. Anti-virus

companies routinely gather over 100,000 unique samples of malicious code in a day, and still many cyber attacks pass undetected.[15] The most advanced attacks, called "zero-day" exploits, epitomize this challenge; such attacks are almost impossible to defend because they use a novel attack method for which there is no signature. Thus, security experts today are forced to defend against broad categories of cyber attacks instead of focusing on individual threats, because it is hard to say exactly what the next cyber attack will look like.[16] The wide variety of possible attack vectors means that a cyber command and control system that restricted use authority narrowly to the topmost national command authority would likely fail catastrophically; by the time policymakers figured out what was happening and how they wished to respond, the damage would be done and the attack might have morphed to new and unanticipated forms, leaving policymakers always several steps behind. Of course, the near-inevitability of surprise could mean that policymakers will be hard pressed to develop the carefully prescribed predelegation conditions of the nuclear era. Therefore, predelegation in the cyber domain may need to be more permissive and flexible than that employed for nuclear command and control purposes.

Most cyber investigations end at a hacked, abandoned computer, after which the trail goes cold.

Third, like nuclear war, cyberwar involves highly technical considerations that even dedicated policymakers are unlikely to master. The cyber sophistication of political leadership can improve with participation in cyber exercises and deeper familiarization with the cyber command and control system. But the rapid evolution of information technology makes it a challenge even for technical professionals to keep pace, so there will likely always be a gulf in understanding between the operators and policymakers. Whereas an inability to understand the finer points of aerodynamics may not limit the quality of political guidance regarding air strikes, confusion over the nature of computer hacking could materially degrade decision-making on cyber response. In a 2010 Black Hat conference keynote address, former CIA Director Michael Hayden stated that conventional operations such as air strikes are discrete events that can be easier than cyber attacks for decisionmakers to manage; the President, he argued, could choose to bomb a factory at any time, but sophisticated cyberattacks take months if not years of painstaking, multifaceted technical subversion. Cyber predelegation, which would allow policymakers to develop guidance focused on desired outcomes in a deliberate manner and well in advance of a crisis, may be the best way for political authorities to get the results they want.

3.2 The Cons: Why Not Cyber Predelegation

Cyber predelegation involves many of the same risks that policymakers wrestled with in the nuclear era. Predelegation would require trusting the cyber operators with decisions that political leaders might prefer to retain for themselves. With cyber, the level to which authority would need to be delegated would likely be even lower in the chain of command than was needed with nuclear predelegation. The complexity and uncertainty of cyber means that predelegation procedures could be especially fraught—specifying in advance the conditions under which certain actions could or could not be taken might be very difficult.

Moreover, the cyber-nuclear analogy breaks down in two ways that cut against the desirability of predelegation. First, the attribution problem is much more acute in the cyber domain than in the Cold War nuclear domain. The most vexing challenge for cyber defense today is that of the anonymous hacker. Smart hackers hide within the international, maze-like architecture of the Internet, leaving a tenuous trail of evidence that often runs through countries with which a victim's government has poor diplomatic relations or no law enforcement cooperation. Most cyber investigations end at a hacked, abandoned computer, after which the trail goes cold. Moonlight Maze, a multi-year investigation which sought to find a hacker group that had successfully stolen U.S. technical research, encryption techniques, and war-planning data, discovered "disturbingly few clues" about its true origin (Adams, 2001).

Vint Cerf, one of the Internet's inventors, recently acknowledged that security was not an important consideration in the Internet's original design. If given the chance to start over, "I would have put a much stronger focus on authenticity or authentication" (Menn, 2011). From a technical perspective, it is theoretically possible to solve the attribution problem. For

example, in the near future, the language of computer networks will be Internet Protocol version 6 (IPv6), which will raise the number of computer addresses from 4 billion to a practically infinite number. Everyone and everything—including each person's individual actions—could be tagged with a permanently associated number. IPv6 also supports (but does not require) Internet Protocol Security (IPSec), which can be used to authenticate Internet traffic. In 2006, Chinese Internet Society chairwoman Hu Qiheng stated that "there is now anonymity for criminals on the Internet in China … with the China Next Generation Internet project, we will give everyone a unique identity on the Internet" (Crampton, 2006).

However, the future of cyber attribution, even in a next-generation network environment, is far from certain. Technologies such as IPv6 may be used to mitigate the threat of anonymous cyber attacks, but human rights groups fear that governments will use this new capability to quash political dissent by reducing online privacy. In 2012, the South Korean Constitutional Court overturned a five-year-old law that required citizens to use their real names while surfing the Web, stating that the rule amounted to "prior censorship" that violated privacy, was technically difficult to enforce, and generally ineffective (Ramstad, 2012). Although it is possible to redress some of the Internet's current technical shortcomings, it is likely that connectivity will continue to outdistance security for many years to come. Progress in attribution will be incremental, involve a slow harmonization of national cyber crime laws, improved cyber defense methods, and a greater political will to share evidence and intelligence.

For the time being, however, the attribution problem will likely limit cyber predelegation to a defensive role. In the absence of reliable intelligence regarding a hacker's true identity, it is difficult to deter, prosecute, and/or retaliate against anyone. For example, in 2008, the U.S. military experienced its most serious cyberattack ever (Lynn, 2010) when malicious code was discovered on U.S. Central Command (CENTCOM) unclassified, classified, and command and control systems. The attack was presumed to be directed by a foreign intelligence agency—perhaps Russia—but the true culprit could not be determined with precision (Shachtman, 2010). However, the Pentagon was forced to undertake a large-scale response to the attack, codenamed Operation Buckshot Yankee. Because the initial attack vector had been the insertion of a removable USB flash drive into a U.S. military laptop in the Middle East, the Pentagon decided to issue a blanket prohibition on the use of flash drives throughout the world (Nakashima, 2010).

The second way in which the nuclear analogy breaks down concerns impact. Cyber attacks, in the extreme, could reach catastrophic levels, but likely not levels contemplated at the middle-range, let alone the extreme range, envisioned in global thermonuclear war. There have been some alarming real-world examples, but many credible national security thinkers are still skeptical of the risk posed by cyber warfare.[17] Nuclear predelegation involved extreme scenarios that were unlikely—and, indeed, never came to pass—and yet whose consequences were so dire that political leaders saw predelegation as an acceptable hedge. Cyber would involve scenarios that were comparatively more likely—indeed, may already have happened—and yet whose consequences were not (yet) seen as so daunting that we should run the risks of predelegation.

Moreover, some of the consequences of predelegation might be readily felt, or at least perceived, in the civilian and political worlds through a loss of privacy and the politically-sensitive blurring of civilian-military divides. Properly circumscribing any predelegated cyber authority would require common agreement on the likely threats, but cyber risk analysis and cyber damage assessments are notoriously difficult and time-consuming endeavors. To date, there is still no legislation in

place that requires U.S. commercial enterprises to employ "best practices" in cyber defense. Many organizations today do not even have a good map of their own network infrastructure, let alone confidence in their network security. In stark contrast to a nuclear explosion, some major cyber attacks go absolutely unnoticed by the public, with only the direct participants witting (Libicki, 2009). If and when a real cyberwar takes place, the attacker's identity should be clear because there will be other, circumstantial evidence,[18] but the often intangible nature of most cyber attacks is likely to make cyber predelegation difficult for national security decision-makers to approve.

In sum, if the odds of a catastrophic cyber attack are low, the consequences perceived to be manageable, and the National Command Authority assumed to be available to manage any future cyber crisis, the political stars may not align quite so readily to pave the way for cyber predelegation.

the political stars may not align to pave the way for cyber predelegation

4. The Cyber Predelegation Sweet Spot?

No analogy works in all respects, but nuclear predelegation holds at least one clear lesson for cyber conflict: if cyber commanders do receive predelegation authority, it will likely be for *defensive* rather than *offensive* operations. In fact, defensive predelegation may be all that is needed—and may even be more than is necessary to confront many cyber threats.

In stark contrast to a nuclear attack, most cyber attacks can be stopped—at least in a tactical sense—with purely defensive measures. There is no immediate need to know who the perpetrators are, where they are located, or their true intentions. The urgency stems from a need to locate, isolate, and neutralize malicious code as fast as possible. Furthermore, blocking malicious data is far easier than shooting down a ballistic missile. In this light, cyber predelegation may not even be necessary, because system administrators already have the authority and capability to protect their networks from what has become an incessant barrage of malware.

Some cyber threats, such as botnets, pose more complicated challenges, and may require cyber defenders to go "outside the wire." Botnet mitigation can even entail the shutdown or hostile takeover of the botnet command and control server(s). But this type of intricate cyber operation, which normally involves the collection of evidence and acquisition of court orders, is unlikely to occur in real-time. To some degree, this seems to obviate the need for cyber predelegation. For example, the celebrated "Coreflood takedown" in 2011 required both Department of Justice (DOJ) user notification and FBI user authorization before the federal government could remove malware from any infected computer (Keizer, 2011).

Still, there may be scenarios in which 1) cyber commanders desire offensive or counterstrike options, and 2) there is simply no time to consult with a traditional chain-of-command. One could imagine that a fleeting window of opportunity would close during which crucial cyber evidence and intelligence could be gained. Here, cyber predelegation might be useful, but its parameters must be governed by the existing laws of war. For example, U.S. forces in Afghanistan are authorized to return fire and pursue adversaries. Logically, cyber predelegation should reflect these same principles. One limitation could be that, in hot pursuit, the counterstrike (or perhaps even a preemptive attack) could not deny, disrupt, degrade, or destroy adversary data or computer resources except in the event that there is no other way to stop a grievous cyber attack on the United States.

Tactical cyber commanders are likely to have rules of engagement that are much more liberal than those given to nuclear commanders, due to the fact that cyber attacks are simply not as dangerous as nuclear attacks. If malicious code is found already installed on a compromised U.S. government computer, defensive actions may be straightforward, as in Operation Buckshot Yankee. If a cyber attack emanates from the U.S. private sector,[19] FBI and DHS could take the lead, with technical support from NSA and CYBERCOM if necessary. When a cyber attack on the United States emanates from a foreign network, it is preferable to contact the country in question's national law enforcement and system administration personnel to help stop it. However, there will be occasions when foreign cooperation is not forthcoming or when there is no time for consultation before costly, irreparable harm would be done to the United States. In this case, predelegation might

authorize preemption or a counterattack against the offending computer or computers.

Due to the attribution problem, this predelegation policy should recognize that U.S. computer networks must be protected even when the assailant is unknown. Positive cyber attribution should be required for significant retaliation, but simple, defensive blocking actions against an ongoing cyber attack will be OK. ICBMs have a return address, but we may never know the true source of some cyber attacks—or we may be successfully deceived by a "false flag" operation. However, even without knowing the true identity of an attacker, CYBERCOM may still be able to target the proximate source of the attack according to the laws of war—with discrimination, proportionality, etc. (Schmitt, 2013). For some forms of cyber attack, such as a denial-of-service,[20] the easiest and most passive form of defense is to "blackhole," or silently discard, the malicious traffic somewhere on the Internet, before it reaches its target.[21] But for the most serious forms of cyber attack, such as a malicious manipulation of U.S. critical infrastructure, CYBERCOM may be able to conduct a pinpoint cyber strike to terminate the malicious process(es) active on the attacking computer, while leaving the other processes intact (if they are presumed to be legitimate).

If neither of these options is possible, the attacking computer may be completely shut down via cyber attack or, in extreme cases, a kinetic attack. This is not ideal, because the attacking computer may have other legitimate processes or functions that could be associated with the national critical infrastructure of another country. Just as soldiers sometimes fire from within hospitals against the laws of war, cyber attackers can also launch attacks from Internet servers that are related to public health and safety. Here, CYBERCOM would have to calculate risk versus reward, but it will need to minimize collateral damage to the extent possible. Any predelegated cyber response should be conducted in legitimate self-defense and supported by as much public transparency as security and intelligence constraints allow. During the operation, CYBERCOM could notify the targeted computer's system administrator and national law enforcement of its actions and their rationale. Today, there are even recommendations that cybersecurity should be included in international security fora at the highest levels of government, with the aid of a mechanism such as the Moscow-Washington "hotline," which is designed to help world leaders defuse international crises (Segal, 2012).

5. Conclusion

The history of nuclear predelegation offers helpful insights into whether and how we should grant predelegation in the cyber domain. Nuclear predelegation was an easy-to-implement workaround that seemed to avoid the potential pitfalls of presidential succession and command devolution. In a similar fashion, cyber predelegation may help USCYBERCOM to defend U.S. critical infrastructure in the new and fast-evolving domain of cyberspace, which, like the nuclear domain, presents vexing challenges to reliable command and control.

There are several similarities between nuclear attacks and cyber attacks, including speed, surprise, and specialization. Together, these characteristics could make some level of cyber predelegation inevitable. However, there are also important differences between nuclear and cyber, including impact and attribution, which USCYBERCOM must consider before granting any level of cyber predelegation.

Unlike a nuclear holocaust, cyber attacks do not pose an apocalyptic threat to the United States, at least not yet. Therefore, they do not pose the always-never dilemma, nor do they demand predelegation. Although attackers have a considerable tactical advantage on the cyber battlefield, it is not clear that they possess a long-term, strategic advantage (Gerth &

Risen, 1999).[22] When the element of surprise is gone, and especially if positive attribution is made, traditional military and diplomatic might should determine the victor in a real-world conflict, and that fact already provides some degree of cyber attack deterrence.

The tactical advantages that hackers enjoy, however, must be addressed, and a national dialogue on cyber predelegation may be the right opportunity. The Internet is worth protecting—it offers a higher level of efficiency, transparency, accountability, and responsibility in government, civil society, and the marketplace. The public would likely support a national effort to give cyber defenders clear rules of engagement, and which also notifies malicious actors of red lines that they may not cross.

As with nuclear predelegation, a stronger case can be made for the use of defensive cyber weapons, especially if their impact is limited to U.S. networks. However, we now know that predelegation during the Cold War extended beyond the use of nuclear weapons in a defensive role, so it is possible that this will happen with cyber weapons as well.

In the nuclear domain, civilian leaders have always demanded that they retain positive control over nuclear weapons. In cyberspace, it is likely that even the public will want to have a say, because civilians now live in this domain. President Eisenhower understood that any nuclear war might take place over allied territory; there is an easy analogy to cyber here: the cyber battles of the future will take place on the same terrain that everyone uses for banking, watching the news, and communicating with their friends and families.

Some aspects of nuclear predelegation and cyber predelegation are similar—how far down the chain of command to go, and how much latitude commanders should have for interpretation. But some characteristics of cyber conflict are unique. Information technology convergence (Dawson, 2003) now sends practically all communications through the same wires, so friendly fire and collateral damage during cyber conflicts may be difficult to avoid. If any cyber attack, even in self-defense, leads to the disruption of Internet sites related to public health and safety, war crimes charges could follow. Finally, information technology is evolving so rapidly that rules for cyber predelegation granted today may not be valid tomorrow.

In summary, political leaders may be forced to authorize some level of predelegation to the military so that it can defend U.S. national sovereignty in cyberspace, but they are also likely to be every bit as skittish about its associated risks. At a minimum, they will want to preserve most of the form and substance of political control. ❇

REFERENCES

Adams, J. (2001) "Virtual Defense," *Foreign Affairs* 80(3) 98–112.

Blair, B. (1985). *Strategic Command and Control: Redefining the Nuclear Threat* (Brookings).

Bracken, P. (1983) *The Command and Control of Nuclear Forces* (Yale Press).

Broad, W.J., Markoff, J. & Sanger, D.E. (15 Jan 2011) "Israeli Test on Worm Called Crucial in Iran Nuclear Delay," *New York Times*.

Burr, William. (12 December 2012) "U.S. Had Plans for 'Full Nuclear Response' In Event President Killed or Disappeared in an Attack on the United States," posted on the National Security Archive. Available here: http://www.gwu.edu/~nsarchiv/nukevault/ebb406/

Crampton, T. (19 March 2006) "Innovation may lower Net users' privacy," *The New York Times*.

Dawson, R. (2003) *Living Networks: Leading Your Company, Customers, and Partners in the Hyper-Connected Economy,* Ch.7: "The Flow Economy: Opportunities and Risks in the New Convergence," (New Jersey: Prentice Hall) 1418–168.

Feaver, P. (1992) *Guarding the Guardians: Civilian Control of Nuclear Weapons in the United States* (Cornell Press).

Gerth, J. & Risen, J. (2 May 1999) "1998 Report Told of Lab Breaches and China Threat," *The New York Times*.

Gorman, S. (7 May 2009) "FAA's Air-Traffic Networks Breached by Hackers," *The Wall Street Journal*.

Hoffman, D. (2010) *The Dead Hand: The Untold Story of the Cold War Arms Race and Its Dangerous Legacy* (Anchor).

Kaplan, F. (1983). *The Wizards of Armegeddon* (Simon & Schuster).

Keizer, G. (27 APR 11). "Feds to remotely uninstall Coreflood bot from some PCs." *Computerworld*.

Libicki, M.C. (2009) "Sub Rosa Cyber War," *The Virtual Battlefield: Perspectives on Cyber Warfare*, Czosseck, C. & K. Geers (Eds) (Amsterdam: IOS Press) 538–65.

Lieber, K. "The New History of World War I and What it Means for International Relations Theory," *International Security* 32, no. 2 (Fall 2007), pp. 1558–191.

Lynn, W.J. (2010) "Defending a New Domain: The Pentagon's Cyberstrategy," *Foreign Affairs* 89(5) 978–108.

Menn, J. (11 October 2011) "Founding father wants secure 'Internet 2'," *The Financial Times*.

Meserve, J. (26 September 2007) "Sources: Staged cyber attack reveals vulnerability in power grid," Cable News Network.

Nakashima, E. (25 August 2010) "Defense official discloses cyberattack," *The Washington Post*.

Orr, R. (2 August 2007) "Computer voting machines on trial," Knight Ridder Tribune Business News.

Preimesberger, C. (2006) "Plugging Holes," *eWeek* 23(35) 22.

Rosenbaum, R. *Slate* (9 January 2009). "The Letter of Last Resort".

Ramstad, E. (24 August 2012) "South Korea Court Knocks Down Online Real-Name Rule," *Wall Street Journal*.

Sagan, S. (1989). *Moving Targets: Nuclear Strategy and National Security* (Princeton Press).

Segal, A. (19 December 2012) "Do the United States and China Need a Cybersecurity Hotline?" *The Atlantic*.

Shachtman, N. (25 August 2010) "Insiders Doubt 2008 Pentagon Hack Was Foreign Spy Attack (Updated)," *WIRED*.

Schmitt, M. (Ed.) (2013) *Tallinn Manual on the International Law Applicable to Cyber Warfare* (Cambridge University Press, 2013).

Van Creveld, M. (1985) *Command in War* (Harvard Press).

Wagner, D. (9 May 2010) "White House sees no cyber attack on Wall Street," Associated Press.

NOTES

1 "Instructions for the Expenditure of Nuclear Weapons in Accordance with the Presidential Authorization Dated May 22, 1957,"declassified on April 4, 2001, accessible at the National Security Archive, http://www.gwu.edu/~nsarchiv/NSAEBB/NSAEBB45

2 This section is adapted from Peter Feaver (1992), pp. 38–66.

3 This authority is spelled out in sec. 301 of title 3, United States Code.

4 The first tranche of 16 documents was declassified and published in 1998 and is summarized here: http://www.gwu.edu/~nsarchiv/news/19980319.htm The original declassified documents are available here: http://www.gwu.edu/~nsarchiv/news/predelegation/predel.htm. See also: http://www.gwu.edu/~nsarchiv/nukevault/ebb332/index.htm For a good summary of more recently declassified documents, see Burr (2012). Available here: http://www.gwu.edu/~nsarchiv/nukevault/ebb406/ See also: "An interview with Carl Kaysen," by Marc Trachtenberg, David Rosenberg and Stephen Van Evera, MIT Security Studies Program, available here: http://web.mit.edu/SSP/publications/working_papers/Kaysen%20working%20paper.pdf.

5 Notes of the President's Meeting, October 14, 1968. Available here: http://www.gwu.edu/~nsarchiv/nukevault/ebb406/docs/Doc%205A%20Furtherance%20document%20Oct%201968.pdf

6 As a 1978 Defense Science Board study put it, if the attack came while the President was in Washington, D.C. "it would be possible … for the President either to command the forces until the attack hit Washington and he was killed or to try to escape and survive, but not both." Quoted in "Joint Chiefs of Staff, Joint Secretariat, Historical Division, Joint Chiefs of Staff Special Historical Study, A Historical Study of Strategic Connectivity, 1950-1981 July 1982." Available here: http://www.gwu.edu/~nsarchiv/nukevault/ebb403/docs/Doc%201%20-%20connectivity%20study%201982.pdf

7 A Permissive Action Link (PAL) is a security device for nuclear weapons, whose purpose is to prevent unauthorized arming or detonation of the nuclear weapon.

8 Lieber (2007) argues that the new historiography on World War I casts doubt on the "automaticity" of the mobilization plans. On the Soviet Dead Hand system, which provided for a nuclear response if the system detected a physical signs of a nuclear strike, see Hoffman (2010).

9 The British resolved this public/private question with a "Letter of Last Resort," a hand-written note from the Prime Minister to a submarine commander, normally locked in a safe and presumably never read and then destroyed upon completion of a tour, that provides instructions for what to do in the event of an actual nuclear war.

10 "Memorandum of Conference with the President, June 27, 1958," dated June 30, 1958. Declassified on April 4, 2001, accessible at the National Security Archive, http://www.gwu.edu/~nsarchiv/NSAEBB/NSAEBB45

11 After the Dow Jones surprisingly plunged almost 1,000 points, White House adviser John Brennan stated that officials had considered but found no evidence of a malicious cyber attack.

12 In 2007, California government officials held a hearing on the security of its touch-screen voting machines, in which a Red Team leader testified that the voting system was vulnerable to attack.

13 In 2006, the Sandia National Laboratories Red Team conducted a network vulnerability assessment of U.S. water distribution plants.

14 Department of Homeland Security (DHS) officials briefed CNN that Idaho National Laboratory (INL) researchers had hacked into a replica of a power plant's control system and changed the operating cycle of a generator, causing it to self-destruct.

15 Author interview with Mikko Hyppönen, Chief Research Officer for F-Secure, 11 Nov 2011.

16 For example, there are myriad types of SQL injection, which are impossible to predict individually and are best defended conceptually.

17 Persuasive skeptics include Cambridge University Professor Ross Anderson, *Wired* "Threat Level" Editor and former hacker Kevin Poulsen, *Foreign Policy* editor Evgeny Morozov, cryptographer Bruce Schneier, and even the man who wrote "Cyber War is Coming" in 1993, Naval Postgraduate School Professor John Arquilla.

18 This was the case in Estonia in 2007, for example, when even chocolate shipments to Russia were cancelled.

19 This does not mean that the U.S. organization is ultimately responsible for the attack; rather, a hacker may be using a compromised computer on the organization's network from which to launch the attack.

20 This is an attempt to make a computer or network resource unavailable to its intended users, usually by sending it so much bogus traffic that it cannot respond to legitimate requests.

21 For many networks, this can be done easily enough with a configuration change at an organization's external router, disposing of the unwanted network traffic.

22 Persistent cyber espionage may be an exception to this rule—by 1999, the U.S. Energy Department had determined that cyber attacks from abroad, particularly from China, posed an "acute" intelligence threat to U.S. nuclear weapons laboratories. As stated earlier, all things nuclear may have a strategic quality.

Comparing Airpower and Cyberpower

Dr. Gregory Rattray

SUMMARY

This paper examines the similarities and differences between cyberpower in the modern era and airpower from 1914 to 1945, so that military planners may learn useful historical lessons and gain illustrative insight regarding strategic cyber warfare. It details how the dawning awareness of a new form of strategic warfare has been accompanied by bold assertions about its significance and how the United States should respond. Yet we have only begun to comprehend the underlying considerations that will shape the nature of strategic cyber warfare. Given the rapid pace of technological change and organizational complexity posed by this environment, actors who sustain their capacity to learn and adapt will be the best equipped to address cyber warfare.

OVERVIEW

Applying the Airpower Narrative to Cyber

Although it is a uniquely manmade warfighting domain, cyberspace lends itself to numerous comparisons to other domains, defense developments, and major events and periods in military history. This paper examines the similarities and differences between cyberpower and airpower, so that military planners may learn useful historical lessons and gain illustrative insight regarding strategic cyber warfare. Airpower, defined as the ability to project power from the air and space to influence the behavior of people or the course of events, is illustrative when used to understand strategic characteristics of the cyber environment. During the period between World Wars I and II, particularly, airpower was characterized by emerging technologies and credited with a capacity for enabling quick military victory. As is now the case with cyberspace, airpower allowed significant freedom of movement for military forces and was viewed as an offense-dominant mode of conflict.

air-cyber comparison can show how preexisting assumptions impact military doctrine, policy, and strategy

Given these observations, the air-cyber comparison can show how preexisting assumptions impact military doctrine, policy, and strategy. For air warfare, early doctrinal assumptions about aircraft technology and offense dominance led to an emphasis on strategic bombing, at the expense of recognizing other important developments that would impact airpower's use in future conflicts. For example, prior to World War II, existing doctrine led the U.S. to overlook lessons from exercise outcomes and others' conflicts which showed bombing forces would not always reach and destroy their intended target. Technological advances which improved the speed of pursuit aircraft and radar detection were not sufficiently addressed, causing setbacks for U.S. bombing campaigns in World War II.

While characteristics unique to cyberspace may allow the U.S. to avoid similar problems in preparing for future conflict, the air-cyber comparison highlights the potential adverse consequences of such doctrinal "lock-in," and the importance of flexibility in strategic thinking.

To illustrate these points, this paper does not solely focus on lessons from the interwar period. This paper will supplement its analysis by referencing environments and examining the lessons from military engagements beyond the interwar period. It will build on the author's previous works, *Strategic Warfare in Cyberspace*, and "An Environmental Approach to Understanding Cyberpower" [Rattray 2001], a chapter in the National Defense University's 2010 *Cyberpower and National Security* [Rattray 2010].

WHAT ASPECTS OF THE CYBER ENVIRONMENT DOES THE AIR ENVIRONMENT HIGHLIGHT?

General Characteristics of the Cyber Environment; Comparison to Air Environment

Cyberspace comprises both the physical and logical systems and infrastructures that are governed by the laws of physics and the logic of computer code. The principle physical laws governing cyberspace are those related to electromagnetism and light. The speed at which waves propagate and at which electrons move creates both advantages and challenges: global communications across cyberspace can happen nearly instantaneously, and vast amounts of data can rapidly transit great distances, often unimpeded by physical barriers and political boundaries. Cyberspace's speed and freedom of movement create both challenges and advantages for individuals, organizations, and states, but at the same time, create weaknesses that could be exploited by adversaries.

In prior analysis, the author has compared elements of this environment to the land, sea, air, and space domains, based on the following four characteristics: technological advances, speed and scope of operations, control of key features, and national mobilization [Rattray 2010]. In the air environment, technological advances eventually allowed attacking forces to strike at centers of gravity directly. For cyber, technology creates new strategic vulnerabilities, and empowers non-state actors. With high speed and broad scope of operations in the air environment, conflicts can end quickly, while cyber is even faster, with the automation of command and control. In the air, the control of key features makes first strikes against adversary airfields crucial, while in cyber, key features are under human control in a highly malleable environment. Given the integral role of changing dual-use technology in both the air and cyber environments, it is important that a government maintain a cadre of professionals and engage with the private sector for the purpose of national mobilization [Rattray 2010].

National security organizations thus cannot simply defend the cyber environment by increasing the size of their military cyber forces. If the attacker has a high probability of rapid success, simply pursuing current cyber defense approaches with more vigor is unpromising. Most attention in the national security community has focused on risks from cyber espionage or a single, time-limited strategic blow from a major adversary. Counterstrategies to deal with state or terrorist non-state actors conducting an economic guerrilla campaign in cyberspace are still not fully developed. A robust, defensible infrastructure will depend on shaping the technologies employed, the obligations of operators of key networks and infrastructures, and the ability to coordinate government-private sector investment and responses to attacks. Distinct from the air environment, key features in the cyber environment require collaborative efforts between the public and private sectors.

Offense Dominance as an Important Characteristic in Air and Cyberspace

Like the air environment, offense dominance is characteristic of cyberspace. However, this has very different implications in cyberspace. Both the weaknesses in the security features of the technological foundations and the economic incentives for openness between networks and systems have made much of cyberspace vulnerable to exploitation, manipulation, and disruption by digital attack, especially networks operated by commercial enterprises,. Non-state actors derive advantages from the ability to leverage expertise, and make decisions rapidly. Generally, offense is easy, defense is difficult.

Concerns over which actor might strike first in a conflict play out differently in cyberspace than with air forces or the use of ballistic missiles. The ease of stealthy deployment of attacking forces and difficulty in attributing the source and intent of attackers mean that damage limitation through preemptive first strikes or retaliatory strikes is largely irrelevant: an actor would have little confidence in trying to attack preemptively to remove the cyber attack forces of an even moderately sophisticated adversary. Similarly, trying to use cyber counterattack to thwart attacks in progress is complicated by issues of identifying and discretely targeting a complex web of electronic points of origin of the attacker. The culpability of network owners and systems from which attacks appear to originate, and the fundamental fact that disrupting these points in cyberspace may only have a limited effect, are further complications.

Deterrence by retaliation is problematic, given the difficulty of attributing an attack to an identifiable perpetrator.

Aspects of the Mission: Learning from the Battle of Saint-Mihiel

The Battle of Saint-Mihiel was the first mass operation of airpower during wartime, when in 1918 then-Colonel William Mitchell assembled U.S. aircraft to support allied ground troops under General John Pershing against German forces. This particular use of airpower may illuminate aspects of cyber warfare when conducted in support of conventional forces. This analysis is based on previous work entitled "Categorizing and Understanding Offensive Cyber Capabilities and Their Use," written jointly with Jason Healey and provided for the 2010 National Research Council *Workshop on Deterring Cyberattacks* [Rattray/Healey 2010].

Though there have been information warfare organizations since the mid-1990s, known cases of cyber attacks in support of military operations are limited in number and not a part of large-scale operations. However, in the future, cyber conflict will result in numerous instances where cyber forces engage heavily in support of traditional military operations. Cyber forces would be strongly integrated with kinetic forces, and the scope of the effect of cyber technology would be operational in nature. Unlike an overt force-on-force cyber conflict or cyber "Battle of Britain," as described below, a cyber St. Mihiel would most likely occur in the context of an existing state of war.

cyber forces will engage heavily in support of traditional military operations

Analogizing from the air support efforts of St. Mihiel, a comparable cyber engagement would involve computer network defense of U.S. communications, as well as computer network attacks against the adversary's systems. Computer network attacks could also be used to destroy or delay the transmission of information from the opposing forces, while also targeting command and control. Computer network exploitation could provide intelligence on the adversary's intent and potential movements. Kinetic and cyber attacks might ensure access to key information like troop movements, adversary command decisions and intent, combat assessment for kinetic and cyber strikes, and generate new information about enemy vulnerabilities [Rattray/Healey 2010].

In a cyber St. Mihiel scenario, the adversary may have cyber capabilities to respond in the cyber domain, but at other times it may be that one side has superiority in cyberspace. Either way, deterrence in this environment will not only have failed, but it will be a secondary consideration compared to controlling the kinetic conflict. It may still have some role, however, if patriotic hackers and copycat attacks confuse each side's national leaders and upset conflict resolution processes.

Aspects of the Mission: Learning from the Battle of Britain

The Battle of Britain, the series of German aerial campaigns against the British in 1940, may aid understanding of what an interstate cyber-on-cyber conflict would look like. As with this paper's discussion of the Battle of Saint-Mihiel, this comparison is based on the work written with Jason Healey for the 2010 National Research Council's *Workshop on Deterring Cyberattacks* [Rattray/Healey 2010]. During the first large-scale contest between air forces, the original German goal was to enable an invasion of Britain through attacks on ports and convoys. The Germans then changed focus to the Royal Air Force (RAF), and later directed attacks against key cities through terror bombing. In the end, these attacks meant to inflict strategic damage, either through destroying national morale or through direct economic damage from destroyed infrastructure. The British used defensive engagements combined with their own limited offensive strikes. Because defense was the primary objective of the RAF, its limited offensive strikes targeted barges to thwart invasion, but involved some strategic attacks against Berlin as well [Rattray/Healey 2010].

A cyber equivalent of the Battle of Britain would be fought completely within cyberspace, involving attackers and defenders on both sides. One key difference in this cyberspace scenario is that both sides would use a significant number of private sector attackers and defenders, raising new concerns related to combatant status in cyberspace. While both sides

may continue to escalate violence in cyberspace, this does not necessarily guarantee that escalation will result in a resort to traditional kinetic warfare, given concerns about international reaction and the potential for significantly greater casualties [Rattray/Healey 2010].

As in an extended aerial campaign, success in a cyber Battle of Britain would go to the side with the more flexible employment of tactical and operational forces, especially if that side had a doctrine that most closely matched the conditions of the future battle. The British held a decisive advantage with their "Dowding System"[1] to detect incoming attack formations and respond quickly and effectively. To detect attacks the British developed and employed both high-tech (radar) and low-tech (searchlights and observer corps) systems. The heart of this system was an information control center to gain comprehensive situational awareness based on all the incoming feeds, allowing quick assessments about incoming attacks. This knowledge improved defensive decisions which were passed to operational commanders who then conducted their own interceptions.

A defender in a cyber Battle of Britain, would have directly analogous needs, from equivalent radar (technology like deep-packet inspection at Tier 1 telecommunication providers to sense incoming attacks) to an observer corps (concerned users or companies reporting incidents to the government or information sharing centers). This information would have to pass to a command center empowered to see all the data and with sufficient authority to issue orders. A cyber command center, for a cyber Battle of Britain, would present several tremendous disadvantages compared to the RAF's Fighter Command Headquarters. In a cyber conflict, attacks may target the private sector which may be outside of the military commander's authority. To address this challenge, Tier 1 telecommunications providers, financial institutions, power companies and other private sector targets would need to have a way of sharing situational awareness and coordinating with the military. While the RAF's commands were able to issue orders to subordinate commands, most Western nations lack any ability to issue authoritative orders to critical infrastructure sectors that would be targeted by cyber attacks.

Tier 1 telecommunications providers, financial institutions, power companies and other private sector targets would need to have a way of sharing situational awareness and coordinating with the military.

Aspects of the Mission: The Need for Trained and Ready Forces in Both Domains

Technology's integral role in air warfare and cyber warfare is clear, but its usefulness is significantly diminished if it is not operated by well-trained and available personnel. Without prepared, experienced personnel, a military does not have the organizational capability for strategic warfare in air or cyberspace.

The initial U.S. strategic bombing campaigns in World War II illustrated the challenges presented by the lack of trained and ready personnel necessary to conduct effective operations. The development of technological knowledge within the Army Air Arm leading up to this point clearly followed the impetus created by doctrine and organizational leadership. A crucial problem was the underdevelopment of human expertise. A principle reason for this underdevelopment was the challenge of training large numbers of personnel to perform a whole range of necessary functions. Even prior to the expansion of the late 1930s, Army Air Arm leadership lamented the lack of enough experienced personnel.[2] Recruiting and teaching enough pilots to fly the tens of thousands of planes envisioned in the mobilization plans required the revamping of training procedures and curricula. The Army Air Arm relied heavily on civilian schools for this purpose, and had difficulty acquiring instructors with the right expertise.[3] Establishing training programs for bombardiers, navigators, radio operators, gunners, and ground crews was even more difficult [GHQ Air Force 1938; Walters 1947]. The specifically military tasks in which these personnel engaged lacked any basis in the civilian sector and far fewer qualified personnel to conduct the necessary training existed within the prewar Air Corps. Throughout the mobilization period and during the war, the U.S. suffered from a lack of qualified instructors for bombardiers and navigators and had to rely on RAF schools and observers to provide

lessons based on combat experience. Maintaining the skill base of instructors for teaching aircraft repair and maintenance also proved difficult because these personnel were pulled toward better paying jobs in depots and factories.

Numbers of personnel aside, an innovative spirit within the nation's cyber workforce will also be a crucial resource in the cyber environment, which rewards pioneers. Risks in cyberspace are less physical than they were for previous explorers, so the premium is on brainpower, creativity, and ability to manage complexity. Historical U.S. strengths—advanced education, systems integration, and intellectual property development and management—should offer advantages in cyberspace competition. In an earlier era, commander of GHQ Air Force Frank Andrews implicitly addressed many of these factors in his advocacy for greater U.S. airpower:

> The tactical and strategical employment of Air Forces and the status of development of aeronautical science exercise a profound influence, each upon the other. The needs of employment spur designers and manufacturers to produce equipment that can meet those needs, and likewise, the equipment on hand, or definitely foreseen, limits and extends the sphere of influence of Air Power [Memorandum to the Assistant Secretary of War 1937].

In light of the relationship between technological change and strategic characteristics of the cyber environment, limitations on capability for strategic cyber warfare are not as much a matter of doctrine as they are a matter of organizational technological capability. The demands of the cyber environment have outpaced U.S. organizational change and currently continue to expand beyond U.S. Cyber Command's (CYBERCOM) capacity and manpower. Even though CYBERCOM's five-fold increase in staffing will be a step in the right direction [Nakashima 2013], further capacity-building efforts must seek a workforce with basic technological skill sets for cybersecurity and cyber warfare. Additionally the U.S. must address broader concerns over declining scientific and technological skills among the population. Increased public support for science, technology, engineering, and math (STEM) education can build entry-level skills necessary for a well-prepared cybersecurity workforce [Corrin 2013]. While CYBERCOM now strives to substantially increase its military and civilian staff, and military leaders debate its possible elevation to a unified combatant command, engagement and further development of these efforts is crucial.

As an additional point in discussing developing organizational capabilities in cyberspace, the lack of requirements for major resource investments, and the ease of leveraging global access to networks, provides more advantages to non-state and lesser state competitors in developing significant capabilities than in other environments. Knowledge of the vital characteristics of critical infrastructures, economic flows, military dependencies, operating systems, and disruptive code can be rapidly stored, duplicated, transferred, and acted upon. Such knowledge and network access permit action in cyberspace on a level far greater than that of the air environment, and allows for faster development of appropriate human capital. This distinct characteristic of the cyber environment makes it all the more important that U.S. efforts achieve nimble and well-staffed cyber warfare capabilities.

SIMILARITIES AND DIFFERENCES BETWEEN THE AIR AND CYBER DOMAINS

Similarities

Doctrinal Advocacy for New Technology's Military Potential Appears in Both Periods

The legacy of World War I influenced the airpower theorists of the early and mid-20th century, in particular Giulio Douhet of Italy, William Mitchell of the United States, and Hugh Trenchard of Great Britain. All three were participants in the rapid development of airpower in the Great War, and they drew similar conclusions about its future role in warfare. As the technology of the airplane rapidly improved, it would enhance the capacity of airpower to strike directly at an enemy homeland, "smashing the material and moral resources of a people," said Douhet, "until the final collapse of all social organization" [Douhet 1942]. Trenchard asserted that "the ratio of morale to material effect was 20:1" [Trenchard 1919]. The bomber, he claimed, would dominate the air and be effectively

unstoppable by defenses. "Viewed in its true light, aerial warfare admits no defense, only offense," argued Douhet, failing to anticipate defensive technology such as radar and advanced interceptors. Future wars, argued these three theorists, would be short and dominated by those with sufficient airpower. Large land or sea forces, or extensive mobilization, would be unneeded. Surprise and preemptive airstrikes would constitute the strategic imperative for all advanced nations. According to Mitchell,

> The advent of air power, which can go straight to the vital centers and neutralize or destroy them, has put a completely new complexion on the old system of making war. It is now realized that the hostile main army in the field is a false objective [Mitchell 1930].

The airpower theorists were not particularly concerned with broader issues of grand strategy and national power, although Mitchell stressed the need to make airpower a national priority to ensure the ability to keep up with rapid technological change.[4] Mitchell argued that airpower could provide a cheap source of security and avoid the large expenditures, conscription, and taxes required to maintain standing armies.

Now, as with early stages of airpower, the evolution of cyberspace is enabling new approaches and forms of warfare. The U.S. Department of Defense (DoD) has pushed aggressively over the last two decades toward net-centric operations based on digital communications and ease of access to information at all levels, down to the individual soldier on the battlefield. Special forces units mounted on horseback operating against Taliban positions in Afghanistan called down GPS-guided precision airstrikes from B-52s they could not see [Rumsfeld 2002]. New U.S. fighter aircraft such as the F-22 and F-35 carry sensor systems that allow them to share data in real-time about activity in the electromagnetic spectrum, both with higher headquarters and with other units conducting tactical operations. As early as the 1990s, military experts also viewed cyber attacks as a strategic weapon for direct strikes against an adversary's centers of gravity, based on the potential for such attacks to degrade and disable communications, electric power supply, and finance [Swett 1995; Schmitt 1999].

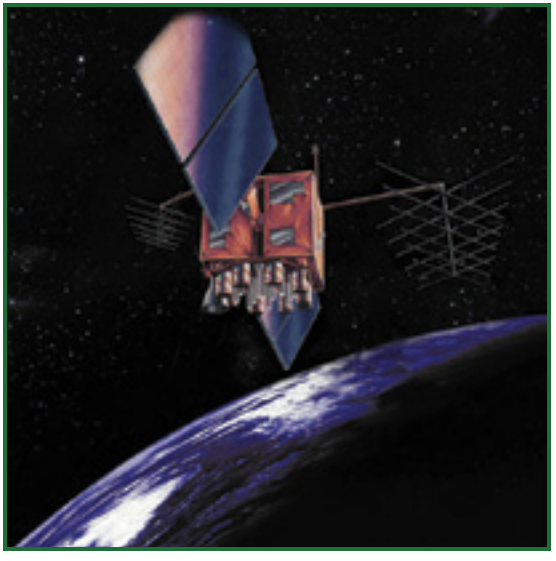
The nature of this new and open environment to facilitate sudden, at-will attack bears much similarity to the air environment of Douhet and Mitchell, as does the belief that cyber warfare is primarily offensive in nature [Garamone 2010; Benitez 2012; Gompert 2011]. Some leading thinkers and policymakers, including former Deputy Secretary of Defense William Lynn, have specifically cited the openness and fluidity of the Internet as factors which favor offense dominance [Lynn 2010]. More recently, the current Director of the National Security Agency and Commander of CYBERCOM, General Keith Alexander has made a similar point, when in 2011 he stressed the need for offensive military cyber capabilities [Talbot 2011].

However, we know from airpower, the dangers of untested doctrine. Because such developing beliefs have not yet been tested in a sustained strategic cyber conflict, and because this form of conflict itself is difficult to define, policy makers must exercise care in developing, establishing, and implementing related doctrine.

Periods of Rapidly Advancing Technological Performance, Short Technology Life Cycles

As with cyberpower today, airpower has experienced periods of rapidly advancing technology and short technology life cycles which in turn rendered doctrinal assumptions about strategic air warfare less useful. The airpower theorists thought that the rise of unstoppable strategic bombers would mean that direct strikes at the enemy centers of gravity would decide future conflicts. Pre-World War II developments such as improved

pursuit aircraft and radar, however, undermined the decisive impact of strategic bombers, placing new demands on airpower strategy. Emphasis on the offensive detrimentally affected defensive considerations and the need to develop pursuit and escort aircraft. Leading up to World War II, lessons learned by the air forces during World War I were gradually forgotten at the Air Corps Tactical School (ACTS) [Greer 1985]. Although thinkers within the Air Corps continued to believe in the need for air superiority, operational concepts in the U.S. came to mirror the earlier thoughts of Douhet and Trenchard concerning the inability of pursuit aviation to intercept and destroy bomber aircraft. Later reflections by bomber advocates at ACTS and planners of the U.S. strategic bombing campaigns lamented the lapse in attention to pursuit aviation, but the offensive doctrine that would dominate the thinking of most Army Air Force leaders in World War II had formed well before the conflict.



The space environment also provides similarities related to the impact of technological change, in that the ability of man to move into space has led theorists such as Colin Gray, Geoffrey Sloan, and Mark Harter to argue that sustained space presence will be an essential enabler of both military operations and control over the global information infrastructure [Gray 1999; Harter 2006]. In this environment, the speed of technological change has been fueled by the increased availability of space-based assets, such as global positioning systems (GPS), satellites, and satellite imagery, to private enterprise. As early as the late 1980s, and leading into the 1990s, this stirred fears that adversaries could use this more widely available technology to detect classified facilities or movements revealing military operations [Richelson 2012; Broad 1987].

This fear has led major players in the international community to reinvest in or build their own alternatives. The potential military applications of GPS helped convince the European Union to create a citizen-oriented system known as Galileo that would not be shut off in times of conflict, unlike the American, Chinese, or Russian equivalents [Higgins 2013]. Technological change in this environment, according to Harter, provides linkages with cyberspace, allowing information to be carried globally, and network warfare operations to leverage space systems [Harter 2006]. Space forces will conduct separate, parallel strategic campaigns with a global reach, such as warning and defending against ballistic missile launches. At the level of grand strategy, in his view, space systems can provide a means to exercise other "instruments of national power (diplomatic, informational, military, and economic) to force an enemy to capitulate" [Harter 2006]. Furthermore, he related uses of space satellites and their orbital locations can also create chokepoints in the cyber world. As with the air environment, technological change in the space environment has created new defensive and offensive opportunities which challenge pre-existing doctrinal assumptions about warfare.

The advent of the Internet, the opportunities for information exchange and social dialogue created by the World Wide Web, and the growing ubiquity of wireless and digital connectivity all have implications for the nature of political, economic, and military interactions. Today, social media, cloud computing, and mobile technology are all undergoing waves of significant change, providing a greater variety of access points to commercial and public services, while widening the attack surface area available to military and non-state actors. As these technologies grow in sophistication and become more intertwined with everyday social functions, it is expected that attacks exploiting these rapidly advancing technologies will increase rapidly as well [Corrin 2013].

The practice of hacking has also experienced its own waves of change. Hackers engaging in demonstrative but relatively apolitical acts such as web defacement were far more common in the 1990s and early 2000s. By contrast, in the present day,

politically motivated hackers are far more common, whether they operate in loose organizations such as Anonymous which target state and prominent commercial institutions, or are non-state actors who act on behalf of states [Mills 2012]. In the latter case, hackers with such political motives have taken over websites and placed confrontational messages and other propaganda, but have also engaged in far more disruptive activities. In the spring of 2007, dissidents with ethnic Russian sympathies organized a disruptive series of cyber attacks that affected the Estonian government, banking, and other sectors [Greenemeier 2007]. More recently, distributed denial-of-service (DDoS) attacks against major U.S. banks occurring in late 2012 through early 2013 have led experts and government officials to assert that the attacks are an Iranian retaliation against Western sanctions and cyber attacks such as Stuxnet [Perlroth 2013].

As such technological change creates new strategic vulnerabilities; it can also render moot strategic observations about cyberspace, and create the need for entirely new ways of thinking. The rise of new approaches to politically motivated hacking and the broader availability of social media have facilitated a convergence between information technology and political protest, enabling events such as the Arab Spring [Waters 2012]. New military capabilities and business enterprises require a conscious balancing of opportunity and risk; this demands an analytic discipline that has not yet developed. The U.S. must learn how to protect its cyberspace presence in a cost-effective fashion. This may involve the development of large offensive forces that "roam the Net" protecting commerce; the orchestration of international accords and norms might be able to limit disruptive activity by states against other states and punish non-state actors; and perhaps a new "cyber Manhattan project" that can establish more secure technological foundations for cyberspace.

in the present day, politically motivated hackers are far more common

The Technology in Both Areas Has Significant Dual-Use Applications

Rapid technological change in the 21st century has created new opportunities and capabilities for militaries and commercial entities alike. Airpower can be viewed in a similar light, as rapid advances in aircraft performance during WWI greatly excited airpower advocates in both military and civilian sectors. However, resources for developing new aircraft and supporting technologies were severely limited. The Air Service consolidated its research, development, and procurement activities during the period at Wright Field near Dayton, Ohio, but also pushed the development of a commercial aviation industry. Billy Mitchell recognized the need for a synergistic relationship between civil and military activities related to the use of airpower. In *Winged Defense,* he stated that "[t]he substantial and continual development of airpower should be based on a sound commercial aviation" [Mitchell 2010]. In a lecture to the Army War College in November 1923, Mason Patrick also called attention to "the intimate relation between the commercial and military air fleets, the readiness with which commercial aircraft can be transformed into military aircraft" and stressed that "therefore, in measuring the air strength of a country due weight must be given to both of these components" [Untitled 1923].

Cyber presents its own range of dual-use applications. The more we improve our capacity for encrypting data and traffic for civilian purposes, the easier it is for members of the armed forces, spies, and criminals to find information on the Internet and communicate anonymously. We use JavaScript for a variety of civilian purposes, such as creating client-side scripts that can interact with the user, but it can also be used to deliver military-grade malware via drive-by download attacks. Even a page of text can be used for guerrilla purposes; for instance, Tamerlan and Dzokhar Tsarnaev, the men responsible for the bombing of the 2013 Boston Marathon, found instructions on how to make a pressure cooker bomb from al Qaeda's online magazine *Inspire* [Ordenez 2013] From a more technical perspective, tools such as pcAnywhere, which allows people to access their computers remotely [Symantec 2013], can also be used to add computers to botnets for the purposes of crime or espionage. Just as the advent of aircraft technology highlighted both military and civilian uses, a variety of actors are realizing cyber technology's own dual-use applications.

Speed and Scope of Operations

Both airpower and cyberpower have seen significant shifts toward faster speed and broader scope of military operations.

For the air theorists, the speed of air operations meant that wars would be over quickly, giving dominant advantages to the party that struck first. As Douhet put it:

> Wars will begin in the air, and…large-scale aerial actions will be carried out even before the declaration of war, because everyone will be trying to get the advantage of surprise…for each side will realize the necessity…of ridding the air of aerial means to prevent any possible retaliation [Douhet 1942].

The situation worsened in the nuclear age, as the advent of ballistic missiles armed with nuclear warheads brought decision making timelines down to minutes, while broadening the scale of effects dramatically.

Today, cyber threats can necessitate responses in seconds. Cyberspace can make information on new political developments across the globe available almost instantly. Commercial companies are tightening global supply chains by means of radio-frequency identification systems linked to point-of-sale electronic inventories, increasing efficiencies and lowering costs. Militarily, new forms of rapidly adaptive operations are made possible by use of these systems. Actionable intelligence can be rapidly pushed to cockpits of aircraft or other weapons systems, allowing engagement of high-value targets across very wide areas, as in the U.S. strikes that killed al Qaeda terrorist leader Abu Mus'ab al-Zarqawi in Iraq [Knickmeyer 2006].

More broadly, advanced militaries that can conduct network-centric operations can tightly orchestrate combined arms campaigns, pursuing full-scale combat operations at any time of the day and in any weather, so they can dominate less sophisticated conventional militaries, as the United States did in Operations Enduring Freedom and Iraqi Freedom. However, global connectivity to achieve rapid strategic impact has become a tool for non-state actors as well. Organized criminal activity, Internet posting of beheading videos, and malicious disruption on a global scale can all spread rapidly. Cyberspace provides opportunities for alliances between organized crime, hackers, and terrorists, multiplying the risk to governments, corporations, and other potential targets. The participation of such non-state entities also adds to the potential for confusion in targeting adversaries in cyberspace, as automated decision making processes, when combined with the risk of human error, may lead to inadvertent harm against innocent individuals or government entities.

Given the speed and scope of operations in cyberspace, management and acquisition processes will need to support agility in the adoption of rules governing access to outside networks and mission partners that balance usability and security. The conduct of military and other operations will place a premium on trusting individuals to understand the changes they see in the cyber tactical environment and adjust the execution of their operations quickly.

Differences

Non-state Actors Can Leverage Cyber Warfare Tools More Easily

Unlike airpower, which is primarily the province of states, cyber warfare tools diffuse much more readily to lesser states and non-state actors. The rapidity of connections offered by modern communications, information systems, and developing cyber weapons technologies creates opportunities for industrialized states, but it can also enable lesser states and non-state actors [Wilson 2009; Lachow 2009]. In addition to organized criminal activity and Internet posting of terrorist propaganda, malicious disruption on a global scale can spread rapidly. Cyberspace provides opportunities for alliances between organized crime rings, hackers, and terrorists, multiplying the risk to governments, corporations, and other potential targets.

Cyber Warfare Offers Greater Opportunities for Asymmetric Strategies

Presence in cyberspace and ease of connectivity create new vulnerabilities to attack. Accessibility and anonymity have produced an environment in which smaller organizations and political actors, especially those who seek to avoid vulnerabilities to retribution in other environments, can achieve a disproportionate increase in capabilities to conduct their operations and disrupt those of adversaries.

The increasing use of the Internet and other aspects of the cyber environment by advanced states to orchestrate the operations of their energy, transportation, and other infrastructures create new strategic vulnerabilities. Disruptive effects on

economic, military, and social activities from sustained power outages or loss of confidence in transportation systems could be more severe, involving physical damage and even casualties. Attacks against digital control systems are technologically feasible [Wilson 2007; ICS-CERT 2012]. Such vulnerabilities provide asymmetrical advantages to non-state actors that are less reliant on such control systems and infrastructures. Less developed states can also use this advantage, as exemplified by acts of industrial espionage [ONCIX 2011] and DDoS attacks against civilian infrastructure [Perlroth 2013].

Today, various non-state actors are already using the Internet in aid of their respective causes. A large number of terrorist groups, such as al Qaeda, the Fuerzas Armadas Revolucionarias de Colombia, and Jemaah Islamiyah already use the Internet to recruit, raise funds, educate people in how to perpetuate attacks, and coordinate. However, while the extent to which insurgents and terrorists have conducted cyber attacks is not yet fully clear, various hacktivist groups have carried out DDoS attacks and breaches as forms of protest. The Anonymous collective has shown that it is capable of paralyzing the websites, and thus the operations, of companies such as PayPal [Martin 2013]. In 1998, Milw0rm, a hacker group comprised of a few teenagers, hacked into the Bhabha Atomic Research Centre in Mumbai to protest the Indian government's nuclear tests [Mehta 1998]. As new and more complex forms of cyber weapons become increasingly available, these attacks may transcend online protests and disruptions, resulting in more significant infrastructure disruption and physical damage.

Cyberspace is Not Controlled by Governments of Sovereign States

The number of actors who play a significant role in cyberspace is also a distinguishing feature. States do not, and cannot, control cyberspace to the same degree as they can control land, sea, and air, or even as they could control cyberspace in the past. For example, during both world wars, the U.S. government took control of the operation of the Nation's predominant telephone provider, American Telephone and Telegraph (AT&T) [Rattray 2001]. That was possible because, at that time, AT&T alone provided almost all of the network hardware and determined the communications rule sets that allowed the telephone system to work (although it did so under close regulation by the government). Now, however, in the U.S. and elsewhere, there are myriad providers of devices, connectivity, and services in loosely woven networks with open standards. Western governments have extreme difficulty in controlling the full spectrum of telecommunications and other activities in cyberspace. In more authoritarian regimes, such as China and Iran, national systems of control over content and Internet access are in place with an attendant ongoing cat-and-mouse game with citizens who seek to work around constraints imposed by the government.

Establishing sovereignty, or deciding on rules to govern the global cyber commons, creates major challenges and growing national security concerns for state actors. With telephone networks, governments had ways to control connectivity beyond their borders. However, over time, non-state actors—corporations, non-governmental organizations, public interest groups—have also become influential; it is not just states that set standards or determine the rules of the road. In many respects, governance in cyberspace resembles the American "Wild West" of the 1870s and 1880s, with limited governmental authority and engagement. Users, whether organizations or individuals, must typically provide for their own security. Theories and approaches for exercising state control, and for leveraging control for national power, have not yet been adequately developed.

As evidenced by the proceedings of the 2012 World Conference on International Telecommunication, held in Dubai under the auspices of the International Telecommunication

Union (ITU), there are competing narratives for how cyberspace should be governed [Kehl 2012]. Given the variety of stakeholder, national and civil society interests which govern the numerous functions of the global Internet, no single vision has taken hold to the exclusion of others. Numerous organizations responsible for Internet governance operate under the multi-stakeholder model, meaning that they incorporate the input of industry, governments, technical experts, and civil society in their decision making processes. By contrast, some nation states believe these organizations (and particularly ICANN) disproportionately favor U.S. national and commercial interests, and that states or state-centered organizations like the ITU should exercise greater control over regulatory and security issues. While this alternative view assumes that multi-stakeholder institutions cannot adequately address such concerns, groups like ICANN are responding by working to address security matters and improved representation of the developing world [ICANN 2012]. The challenges of governing this environment thus transcend borders, allowing technical realities of the cyber environment to continue to pose questions about the appropriate role of government.

Identifying Centers of Gravity is More Difficult in Cyber Warfare

Because cyberspace is unique in that interactions are governed by manmade hardware and software, the "geography" of cyberspace is much more mutable than other environments, thus complicating the task of identifying an adversary's centers of gravity. This is not the case in the air, which is a static physical environment. Portions of cyberspace can be turned on and off with the flick of a switch, and can be created or "moved" by insertion of new coded instructions in a router or switch, thus making it more challenging to identify centers of gravity in cyberspace. Many of the rules of code are subject to frequent change, creating new online environments, as exemplified by the emergence of Web 2.0, the continued proliferation of social media, and the use of cloud-based services. While the constant creation of new features in an online landscape complicates targeting, it also creates a wider attack surface area for adversaries, showing that such online activity does not come without vulnerabilities and disadvantages.

Government Cooperation with the Private Sector is Crucial to Effective Cyber Defense Strategy

The U.S. security community has recognized the significant threat posed by digital attacks to both traditional military operations and the nation's well-being. Yet, until recently, national security doctrine has avoided addressing the fundamental role played by the private sector in how risk and vulnerability in the nation's information infrastructures can be jointly managed and defended. Statements regarding the threat posed to the nation's information infrastructures by cyber attacks lump together a very wide range of threats without adequately distinguishing the relative likelihood and risks posed by different categories. Many assessments have concentrated on what U.S. adversaries could potentially disrupt, with very little attention devoted to understanding underlying political objectives of possible adversaries, the degree to which disruption would cause serious damage, or the management of response and recovery efforts after an initial attack.

In order to better protect cyberspace, the United States should pursue redundancy and diversity in undersea cables, satellites, ground stations, and fiber optic routing in order to minimize vulnerable chokepoints. We can worry less about precise mapping of all known vulnerabilities (which have been a focus of many U.S. federal government efforts, given the constantly morphing cyberspace environment). Public and private sector actors who operate and use cyberspace for key national economic and security purposes should jointly conduct regular scenario analyses and exercises to focus investment and develop strategies to establish a robust cyber infrastructure. Because of the adverse national security impacts of widespread economic harm to the U.S. private sector, intellectual property can serve as a center of gravity for adversaries to exploit through espionage. Joint public-private analyses and exercises should not only examine threats to critical infrastructure, but should also account for the adverse national security impact of intellectual property theft. Enabling private sector institutions to exchange information with the government pertaining to data breaches should be facilitated by assuaging fears of prosecution, civil liability, or regulatory action. Even the potential harm of attacks against critical infrastructure should not overshadow the harm that stems from the steadily growing cost of industrial espionage.

LESSONS OF THE AIR-CYBER ANALOGY

The Lessons of Doctrinal "Lock-In"

The adverse consequences of inflexible adherence to doctrine, in spite of technological change, is a key concern for waging strategic warfare. Unlike strategic air warfare, strategic cyber warfare appears less prone to the adverse impact of doctrinal "lock-in" in the U.S., given the fact that it has done more to acknowledge the important role of defensive measures in cyberspace [Fryer-Biggs 2012]. However, the U.S. experience with airpower sets forth important lessons that today's military cyber strategists should consider when shaping policy and strategy.

The first of these lessons is the need to know when operating concepts should be modified or changed, based on more effective defensive measures and detection technology. For airpower in the early-to-mid-1930s, bombers were capable of flying higher and faster than fighters, and this distinction helped lend credence to the emphasis on strategic bombing and the relative disregard of defensive measures. However, technological advances changed this calculation with numerous defensive innovations such as radar that improved warning, better radios for coordinating defensive responses, and the ability of high performance, heavily-armed interceptors to intercept and destroy larger aircraft. But the commitment to existing concepts of operations from the late 1930s through 1943 blinded the Army Air Arm to significant evidence that bombers would not be able to operate effectively during unescorted daylight operations over Germany. Learning this lesson proved very costly for the aircrews of the Eighth Air Force, and impaired the U.S. strategic bombardment effort.

Given the difficulties that the Army Air Force experienced in World War II, cyber forces should also be prepared to address the development of new defensive capabilities in cyberspace. In light of cybersecurity experts' desire to improve attribution technology and thus solve the challenges of identifying malicious actors in cyberspace, significant advances in this area may necessitate a shift in offensive thinking. Current gradual advances in attribution technology do not provide complete certainty as to the identity of an attacker, but strategy should not remain grounded in assumptions that further change will not occur [Lemos 2013]. Changes in attribution technology will make offensive execution a more complex task, in that forces may need to deceive or disrupt defensive mechanisms more completely than today, before pursuing other intended targets. Failure to address an improved detection capability could lead to setbacks for offensive cyber forces similar to those experienced by the Eighth Air Force during its early bombing campaigns against Germany.

Importance of the Role of Public-Private Partnerships

Cyber technology from the private sector has aided the development of strategic warfare capabilities, as was the case for airpower during the interwar period. Around this time, the Air Service placed substantial emphasis on improving the performance of all types of aircraft. Increasing speed and range were primary concerns. According to an active participant in the interwar technological development of the Army air arm, James Doolittle, the involvement of the Air Services in air races and competitions "was for two purposes: one was research and development and the other was to bring aviation to the American public" [Neufeld 1993]. In announcing U.S. military participation in the air races in 1922, a public release by the Secretaries of War and Navy stated that "[t]he encouragement of an aeronautical industry and of aeronautical activity outside the military forces is considered by every nation developing aeronautics the most economical method for developing air power" [Memorandum 1922]. The flight of Army MB-2 bombers around the world in 1924 and Charles Lindbergh's Atlantic crossing in 1927, as well as experiments with refueling, explored the possibilities for improving the range of aircraft.

With its July 2011 Strategy for Operating in Cyberspace (DSOC), the Pentagon expressly linked itself with private industry. Of the five initiatives in the DSOC, Strategic Initiative 3 advocated partnering with other U.S. federal agencies, as well as the private sector, so as to mitigate cyber risks stemming from reliance on private Internet Service Providers (ISPs) and the global supply chain [DoD 2011]. The initiative identified the Defense Industrial Base Cyber Security and Information

Assurance (DIB CS/IA) program as one such effort that it would seek to strengthen through the build-up of related public-private pilot programs. This led to the 2011 creation of the DIB Cyber Pilot program, which leveraged threat data provided by the government to protect the networks of private sector defense contractors [DoD 2012]. The recent White House Executive Order, issued 12 February 2013, offers similar support by seeking to expand the Enhanced Cybersecurity Services program under the Departments of Homeland Security and Defense [The White House 2013]. While sharing threat indicators between public and private entities is one example of collaboration activity for improving cyberpower, this activity exemplifies the need for the government to improve and further work with the defense industrial base in this and other cooperative cyberpower efforts.

ADVANTAGES OF THE COMPARISON

Given its similarities in terms of speed and scope of operations, airpower aids the understanding of the pace of operations in cyberspace. Strategic air warfare requires well-trained personnel to fully utilize technology, and helps illustrate how the same training and experience are of great importance in strategic cyber warfare. While identifying centers of gravity is difficult in both domains, in cyberspace it is distinctly more difficult. The U.S. strategic bombing campaigns in Germany during World War II helped illustrate potential obstacles to effective targeting. Finally, because of the influence of doctrine and the change of technology in air warfare, the air-cyber comparison makes it easier to understand the risks of ignoring the impact of technology on the relationship between offense and defense.

Airpower Aids Understanding of the Pace of Operations in Cyberspace

The development of airpower in the first half of the 20th century meant that attacks could be launched against strategic centers of gravity in hours. Given the capacity for airpower to project force quickly and directly at an adversary's centers of gravity, the airpower-cyberpower comparison illustrates the strategic importance of quick decision making and rapid response. In cyberspace, key events and disruptive threats can necessitate responses in seconds. National leaders are thus faced with tighter timelines for decisions even as it becomes increasingly imperative to orchestrate action across wider distances more quickly. In light of the need for such rapid decision making in strategic cyber warfare, it is necessary to discuss the appropriate balance between automation and human input for analytical purposes. While automation may prove extremely valuable in saving time, the lack of human involvement in certain decisions may present the risk of improperly attacking the innocent.

Comparison Emphasizes the Need for Constant Training and Experienced Personnel

Even as both environments require implementation of highly advanced technology, the U.S. experience with airpower illustrates how technology cannot provide a strategic advantage without a strong cadre of well-trained personnel. This important lesson is yet another function of cyberspace's trend of rapid technological advancement.

U.S. Army Air Forces entered World War II with an underdeveloped organizational technological capacity for strategic warfare, both in terms of numbers of weapons and personnel and in terms of the breadth of technologies, skills, organizations, and training. U.S. airmen were inattentive to developments that changed how strategic bombardment forces could be used more effectively; even if this had been avoided in doctrinal assumptions and planning, sufficient training for U.S. airmen would have been crucial in implementing such strategic foresight. Instead, strategic bombing campaigns were hampered by an insufficient number of skilled bombardiers, navigators, and maintenance personnel.

After World War II, U.S. pilots were successful against Chinese and Korean forces in the air, even if U.S. airpower overall was not effective against enemy centers of gravity. In Vietnam, as with efforts against the Germans in World War II, strategic attacks proved extremely difficult. After Vietnam, the U.S. stood up advanced training facilities at Nellis Air Force Base in Nevada, which placed significant emphasis on realistic near-combat environment training scenarios, so that Air

Force pilots could be adequately trained even in peacetime. Applying these lessons to training for strategic cyber warfare is crucial.

The centrality of human expertise requires the U.S., like other major actors, to compete globally to create, attract, and retain the human capital needed to construct, utilize, and engage in cyber operations. These personnel must be capable of analyzing the ever-changing opportunities and risks present in the environment, operating and protecting the large enterprises and infrastructures that sustain cyberspace, and performing other tasks ranging from forming new modes of sharing information to developing the capacity for preventing or deterring disruptive attacks. For the U.S. military, the challenge is to nurture a strong cadre of cyber experts, similar to the naval, air, and space expertise that has enabled success in other environments. This requires the vision and will to divert resources from traditional military missions to invest in the core capabilities for cyber operations. Today, given the speed at which technologies for cyber warfare develop, education and training are even more crucial, and may have even more of an impact on organizational technological capacity than flexibility in doctrine. Efforts such as the National Initiative for Cybersecurity Education (NICE) under the National Institute of Standards and Technology (NIST), and the Pentagon's cyber workforce development efforts demonstrate that the U.S. is beginning to develop at least some of the skill sets for sufficient cyber organizational technological capacity.

the challenge is to nurture a strong cadre of cyber experts

Inability to Identify and Target Centers of Gravity and Assess Damage

Although identifying centers of gravity and targeting them are more difficult in cyber warfare, the U.S. experience with strategic bombing in World War II highlights such challenges as well as ways they can be mastered. In World War II, the unanticipated ability of the Germans to limit their vulnerability to strategic bombing posed problems for U.S. planners in identifying and attacking the most significant targets. American planners had expected to "find a taut industrial fabric, striving to sustain a large Nazi war effort" [Hansell 1972]. Yet the fact that the Germans had not conducted a full wartime mobilization until 1942 was not evident to either the British or United States. Similarly, the ability of the Germans to redesign systems to minimize use of items like ball bearings and disperse production in the aircraft industry did not figure into targeting schemes.

The Strategic Bombing Survey would later conclude that "[t]he recuperative powers of Germany were immense; the speed and ingenuity with which they rebuilt and maintained essential war industries in operation clearly surpassed Allied expectations" [USSBC 1976]. In particular, the U.S. effort to identify critical nodes for production of finished war materials, to minimize the number of required targets, proved flawed. Attacking underlying systems would have provided higher payoff. As stated by the Strategic Bombing Survey, "[t]he importance of careful selection of targets for air attacks is emphasized by the German experience. The Germans were far more concerned over attacks on one or more of their basic industries and services … than they were over attacks on their armaments industry or the city areas. The most serious attacks were those which destroyed the industry or service which most indispensably served other industries" [USSBC 1976].

Similar challenges will face those who wage strategic cyber warfare. The complexity of modern information infrastructures will make the effects of large-scale attacks difficult to estimate. The ability of the adversary to recuperate must be analyzed. Strategic cyber planners must evaluate which sectors or systems within an infrastructure constitute centers of gravity with the greatest leverage.

The ability of the U.S. strategic bombing campaign to attack the German vulnerabilities identified in war plans was also constrained by the continuing tug of war regarding available heavy bomber assets. This led to a piecemeal commitment of assets to strategic attacks that had very limited effect and may well have allowed the Germans to respond effectively to the threat of heavily armed but unescorted bomber attacks. The continual shift between target systems—from submarines to aircraft and ball-bearings, to supporting the Normandy invasion, to the eventual concentration on oil and transportation—allowed the Germans considerable latitude for

reconstitution and recovery until the final phase of the conflict. In the context of cyber warfare, available assets capable of conducting cyber attacks may well be able to support both battlefield operations and strategic attacks. If the U.S. engages in a conflict with a significant conventional dimension, plans for waging strategic cyber campaigns should similarly expect a competition for resources and possible diversions of effort.

The U.S. strategic bombing campaign in Germany also suffered from an inability to assess the damage it was inflicting. The Army Air Force lacked an intelligence capacity for conducting its own assessments. The assessments conducted by others, such as the Committee of Operations Analysts and the Joint Intelligence Committee, looked to identify future targets without adequately assessing the available information about the effects of attacks already conducted. Estimates measured bombing campaign progress in terms of the numbers and weight of the attacking force, not the effects on the targeted system. As a result, the combined U.S. and British intelligence estimates on the effects of the streategic bombing campaign proved susceptible to wide miscalculation and deception. Even if intelligence-gathering efforts had been improved, the task of understanding the effects of bombing and the German efforts at recovery and substitution were immense. The U.S. Air Force study on intelligence and Army Air Force Operations in World War II stresses the tendency to overestimate damage based on photoreconnaissance and the difficulties posed by the German program to disperse industrial production [Kreis 2004]. Daniel Pape concludes that "[i]nformation was inadequate to produce reliable macroeconomic analysis, let alone comprehensive microeconomic analysis required for strategic interdiction by precision bombing" [Pape 1996]. Even when a thorough analysis such as the Strategic Bombing Survey for Germany was conducted, the assessments of the report were deemed inconclusive in terms of guiding the conduct of the strategic air campaign against Japan. Those responsible for waging strategic cyber campaigns should pay heed to the difficulty of constructing capacity for damage assessment adequate to waging a new type of warfare. Intelligence organizations with the proper skills and analytic tools must exist if planners desire to adapt and improve their strategic cyber warfare targeting plans as a campaign progresses.

In total, those contemplating waging a strategic cyber warfare campaign will confront at least as many challenges in establishing the enabling conditions for success as faced the planners and leaders of the U.S. strategic bombing effort against Germany in World War II. Although achieving an offensive advantage may prove easier in cyberspace, meeting other conditions will present difficulties requiring attention if such campaigns are to prove effective. The lessons of the past should be kept at the forefront of thinking about how to establish cyber warfare forces and wage strategic warfare.

Comparison Aids Better Understanding of the Fluidity Between Offense and Defense

Consistent with the challenges of identifying and targeting centers of gravity, the technological changes which create such concerns also pose the probability of changes in the balance between offense and defense. When judged against the great certainty with which early air theorists regarded offense and the use of strategic bombing, interwar improvements in pursuit aircraft and development of radar technology illustrate the speed at which revolutionary technological change can occur. Today, cyber technology is changing at a far more rapid pace, illustrating the potential for present-day assumptions about the difficulty of attribution and other defensive activities to be upended by innovation.

This has significant implications for a range of activities related to preparation for strategic cyber warfare, including intelligence collection, requirements for public-private

collaboration in crafting new technology, the education and training of forces, and the overall development of military strategy applicable to cyber conflict. As new cyber weapons are created at an increasingly fast pace, fueling the proliferation of threats targeting governments and commercial enterprises alike, it is not unreasonable to view cyberspace as a disproportionately offensive environment. This may, in turn, lead military planners and other cybersecurity experts to make operational assumptions that emphasize their own offensive or active defensive capabilities, while overlooking the need for innovation and adaptation on the defensive side. Again, given the U.S. Army Air Force experience in World War II, the air-cyber comparison demonstrates the usefulness of a more balanced approach to strategic cyber warfare.

DISADVANTAGES OF THE COMPARISON

There are aspects of strategic cyber warfare that cannot be easily understood through the lens of air-cyber comparison. Given the comparative ease of access with which non-state and lesser state actors can enter cyberspace and develop cyber warfare capabilities, cyberpower is within easy reach of smaller parties. Because cyberspace is comparatively more challenging for identifying centers of gravity, is a more malleable environment, and is more open to a variety of actors, it is harder to assess belligerents' use of cyber attacks through detection systems, unlike the case of airpower.

Only States Can Effectively Operate Sophisticated Airpower, the Same Is Not True for Cyberpower

While airpower is a tool that can only be wielded successfully by governments, the tools of cyberpower are very different. For lesser states and non-state adversaries, developing and wielding airpower comes at a far greater expense, and requires far more organizational capacity. Projecting cyberpower does not place the same demands on those who seek it. In comparison to airpower technology, malicious software and other cyber tools can be developed and distributed fairly cheaply, strategically enabling entities that would not have the size or organizational sophistication to succeed in other domains.

Given this important distinction from airpower, U.S. experiences with strategic air warfare and interstate conflict in other domains do little to illustrate lessons on how to deter and respond to non-state adversaries in cyberspace. Military strategists must account for this distinction, and observe lessons on asymmetric warfare from other domains as appropriate. The airpower-cyberpower comparison serves as a caution against rigid over-reliance on doctrinal assumptions, and emphasizes the strategic offensive and defensive implications of vulnerabilities to vital centers. But lessons on asymmetric warfare drawn from U.S. approaches to counterinsurgency and counterterrorism may prove illuminating in efforts to understand the greater variety of potential adversaries in cyberspace.

Easier to Assess Belligerents' Use of Airpower through Detection Systems; Cyberpower More Difficult

The air-cyber comparison generally helps efforts to understand the pace of operations in cyber warfare, but the risk of misidentifying targets in cyberspace is great. The Department of Defense net-centric warfare concepts, fusing improved sensor and communications systems, enables engagement of targets for air attack that emerge rapidly but offer very limited time periods in which to take action. Balancing the need for speed with the risks of automated responses in military and other operations will prove a growing challenge. Rules of engagement will often call for high-confidence identification of potential targets, but a commander may not fully trust automated systems to make the call regarding weapons employment. The U.S. Navy shoot-down of an Iranian airliner in 1988 by the Aegis air defense system provides a cautionary tale, yet *excessive* caution may also lead to a slowed defensive response in the cyber environment [Miklaszewski 2006]. Cyberspace presents chances to hide or mislead regarding the source of malicious activity. Automated systems can be subverted and turned against their operators or used against third parties.

For non-state actors to exploit this unique aspect of cyberspace, a challenge is to take advantage of rapid, global operations without creating a recognizable signature in cyberspace that would render them vulnerable to retaliation, and thus to deterrence. Non-state actors will seek to make cyberspace a

medium where guerrilla campaigns, orchestrated dispersal, and surreptitious disruption make large land, sea, and air forces fighting decisive battles irrelevant.

CONCLUSION

In this century, the United States has entered an environment filled with opportunity and fraught with challenges. Dawning awareness of a new form of strategic warfare has been accompanied by bold assertions about its significance and how the United States should respond. Yet we have only begun to comprehend the underlying considerations that will shape the nature of strategic cyber warfare.

Development of strategic cyber warfare capabilities requires efforts to identify issues and evaluate uncertainties, not react to individual events and accept simple answers. The analysis of strategic cyber warfare requires a deeper understanding of the linkage between applying force and intended political effects. Past efforts to wage strategic warfare have often failed because of inadequate understanding of the ways that applying force related to the political objectives sought rather than because of the shortcomings of technological tools for inflicting pain. Those contemplating strategic cyber warfare must also address the fundamentally new challenges of a manmade cyber environment largely developed and operated outside the control of national governments. Given the rapid pace of technological change and organizational complexity posed by this environment, actors who sustain their capacity to learn and adapt will be best equipped to address cyber warfare. ❇

WORKS CITED

"A New Approach to Africa," 2012, press release, *ICANN.org.*

"Overall Report—European War," 1976, *U.S. Strategic Bombing Survey (USSBC),* Vol. 2, New York: Garland Publishing.

"Symantec pcAnywhere," 2013, *Symantec.*

Benitez, J. and Jason Healey, 2012, "Cyber Offense is King," *Atlantic Council: New Atlanticist Policy and Analysis Blog.*

Broad, W., 1987, "Private Cameras in Space Stir U.S. Security Fears," *The New York Times,* C1, C12.

Corrin, A., 2013, "Cyber threats and agency costs expected to climb in 2013," *FCW.com.*

Corrin, A., 2013, "Desperately seeking cybersecurity pros," *FCW.com.*

Department of Defense, 2011, "Department of Defense Strategy for Operating in Cyberspace," *Defense.gov.*

Department of Defense, 2012, "DoD Announces the Expansion of the Defense Industrial Base (DIB) Voluntary Cybersecurity Information Sharing Activities," *Defense.gov.*

Douhet, G., 1942, *Command of the Air*, New York: Coward-McCann.

Fryer-Biggs, Z., 2012, "U.S. Military Goes on Cyber Offensive," *Defensenews.com.*

Garamone, J., 2010, "Lynn: Cyberwarfare Extends Scope of Conflict," *Defense.gov.*

GHQ Air Force, Office of the Commanding General, 1938, "Tentative Principles Governing Specialized Training," NARG 18, file no. 247.

Gompert, D. and Phillip Saunders, 2011, *The Paradox of Power,* Washington, DC: National Defense University Press.

Gray, C. and Geoffrey Sloan, 1999, *Geopolitics, Geography, and Strategy,* London: Frank Cass.

Greenemeier, L., 2007, "Estonian Attacks Raise Concern Over Cyber 'Nuclear Winter'," *Informationweek.com.*

Greer, T., 1985, *The Development of Air Doctrine in the Army Air Arm,* Washington, DC: Office of the Air Force History.

Hansell, H., 1972, *Air Plan that Defeated Hitler,* Atlanta, GA: Higgins-McArthur.

Harter, M., 2006, "Ten Propositions Regarding Space Power: The Dawn of a Space Force," *Air and Space Power Journal 20,* Vol. 2.

Higgins, A., 2013, "Europe's Plan for GPS Limps to Crossroads," *NYTimes.com.*

Industrial Control Systems Cyber Emergency Response Team (ICS-CERT), 2012, "ICS-ALERT-12-046-01A-(UPDATE) Increasing Threat to Industrial Control Systems," *ICS-CERT Alert.*

Kehl, D. and Tim Maurer, 2012, "Did the U.N. Internet Governance Summit Actually Accomplish Anything?," *Slate.com.*

Knickmeyer, E. and Jonathan Finer, 2006, "Insurgent Leader Al-Zarqawi Killed in Iraq," *Washington Post.*

Kreis, J., 2004, *Piercing the Fog: Intelligence and Army Air Forces Operations in World War II,* Washington, DC: University Press of the Pacific.

Lachow, I., 2009, "Cyber Terrorism: Menace or Myth?" *Cyberpower and National Security,* Washington DC: National Defense University Press.

Lemos, R., 2013, "More Data on Attackers, But Attribution Still Dodgy," *Darkreading.com.*

Lynn, W., 2010, "Defending a New Domain: The Pentagon's New Cyberstrategy," *Defense.gov.*

Martin, A. and Mark Duell, 2013, "Respectable face of newlywed hack whose cyberattacks paralyzed PayPal," *Mail Online.*

Mehta, A., 1998, "Milworm Bites BARC," *Outlook India.*

Memorandum for public release by the Secretary of War and Secretary of Navy on Pulitzer Trophy race, 1922, Narg 18, file no. 229.

Memorandum to the Assistant Secretary of War, 1937, "Procurement Program for the Air Corps 1940–1945."

Miklaszewski, J. and Kerry Sanders, 2006, "U.S. Passes Up Chance to Strike Taliban: Predator had Suspected Fighters in its Sights, but Military Passed on Shot," *NBC News*.

Mills, E., 2012, "Old-time Hacktivists: Anonymous, you've crossed the line," *CNET.com.*

Mitchell, W., 1930, *Skyways: A Book on Modern Aeronautics,* Philadelphia: JB Lippincott.

Mitchell, W., 2010, *Winged Defense,* The University of Alabama Press.

Nakashima, E., 2013, "Pentagon to boost cybersecurity force," *Washington Post*.

Neufeld, J., 1993, *Research and Development in the U.S. Air Force,* Washington, DC: Center for Air Force History.

Office of the National Counterintelligence Executive (ONCIX), 2011, "Foreign Spies Stealing U.S. Economic Secrets in Cyberspace," *NCIX.gov.*

Ordenez, F. and Greg Gordon, 2013,"Radical magazine 'Inspire' may have motivated other terror aspirants," *McClatchy.*

Pape, R., 1996, *Bombing to Win: Airpower and Coercion in War,* Ithaca, NY: Cornell University Press.

Perlroth, N. and Quentin Hardy, 2013, "Bank Hacking Was the Work of Iranians, Officials Say," *New York Times.*

Rattray, G. and J. Healey, 2010, "Categorizing and Understanding Offensive Cyber Capabilities and Their Use," *Proceedings of a Workshop on Deterring Cyberattacks*.

Rattray, G., 2001, *Strategic Warfare in Cyberspace,* Cambridge, MA: MIT.

Rattray, G., 2010, "An Environmental Approach to Understanding Cyberpower," *Cyberpower and National Security,* page 253.

Richelson, J., 2012, "Declassified Documents Trace U.S. Policy Shifts on Use of Commercial Satellite Imagery from 1970s to Today," *George Washington University National Security Archive.*

Rumsfeld, D., 2002, "Transforming the Military," *Foreign Affairs.*

Schmitt, M., 1999, "Computer Network Attach and the Use of Force in International Law: Thoughts on a Normative Framework," *Technology, Policy, Law, and Ethics of U.S. Cyberattack Capabilities,* 37:885–937.

Swett, C., 1995, "Strategic Assessment: The Internet," *FAS.org.*

Talbot, D., 2011, "Should We Fire the First Shot in a Cyberwar?" *MIT Technology Review.*

The White House, 2013, "Executive Order—Improving Critical Infrastructure," *Whitehouse.gov.*

Trenchard, H., 1919, "Report on the Independent AirForce," 1334–1335.

Untitled lecture, 1923, NARG 18, file no. 229.

Walters, R., 1947, *Individual Training in Aircraft Armament by the AAF 1939-1945*, AAF Historical Offices, Army Air Forces Headquarters.

Waters, T., 2012, "Social Media and the Arab Spring," *Smallwarsjournal.com.*

Wilson, C., 2009, "Cyber Crime," *Cyberpower and National Security,* Washington DC: National Defense University Press.

Wilson, T., 2007, "Experts: U.S. Not Prepared for Cyber Attack," *DarkReading.com.*

NOTES

1 The Battle of Britain: The Dowding System. *Spiritus Temporis*. Available at http://www.spiritus-temporis.com/battle-of-britain/the-dowding-system.html, accessed June 13, 2013.

2 For evidence of such concerns, see lecture by Maj. Gen. Frank Andrews, Commander, GHQ Air Force, to Army War College, "The General Headquarters Air Force," October 9, 1937, 3, in Air Force Historical Research Agency (AFHRA) file no. 415.201; and his lecture to ACTS, "Problems Met by the GHQ Air Force and Solutions to Some of Them," April 1938, 5, in AFHRA file no. 248.2019A-19.

3 See "Conclusions," 112, of Army Air Force Historical Studies, no. 61, "Combat crew and Unit Training in the AAF 1939-1945," in AFHRA file no. 101-61.

4 The need for a national effort is the primary focus of William Mitchell, 1925, *Winged Defense: The Development and Possibilities of Modern Airpower—Economic and Military,* New York: G.P. Putnam's Sons.

Active Cyber Defense: Applying Air Defense to the Cyber Domain[1]

Dorothy E. Denning & Bradley J. Strawser
Naval Postgraduate School

In the domain of cyber defense, the concept of active defense is often taken to mean aggressive actions against the source of an attack. It is given such names as "attack back" and "hack back" and equated with offensive cyber strikes. It is considered dangerous and potentially harmful, in part because the apparent source of an attack may be an innocent party whose computer has been compromised and exploited by the attacker.

Our purpose in writing this paper is to show that active cyber defense is a much richer concept that, when properly understood, is neither offensive nor necessarily dangerous. Our approach is to draw on concepts and examples from air defense to define and analyze cyber defenses. We show that many common cyber defenses, such as intrusion prevention, have active elements, and we examine two case studies that employed active defenses effectively and without harming innocent parties. We examine the ethics of active cyber defenses along four dimensions: scope of effects, degree of cooperation, types of effects, and degree of automation. Throughout, we use analogies from air defense to shed light on the nature of cyber defense and demonstrate that active cyber defense is properly understood as a legitimate form of defense that can be executed according to well-established ethical principles.

We are by no means the first authors to address the ethics of active defense. Dittrich and Himma (2005), for example, contributed substantially to initial thinking in this area. Our work differs from theirs and other work in this area through its application of air defense principles. We believe that the analogy of air defense helps shed light on active cyber defense and the moral issues it raises.

DEFINING ACTIVE AND PASSIVE CYBER DEFENSE

Because our definitions of active and passive cyber defense are derived from those for air defense, we begin by reviewing active and passive air and missile defense.

Active and Passive Air and Missile Defense

Joint Publication 3-01, *Countering Air and Missile Threats*, defines active air and missile defense (AMD) as: "direct defensive action taken to destroy, nullify, or reduce the effectiveness of air and missile threats against friendly forces and assets." The definition goes on to say that active AMD "includes the use of aircraft, AD [air defense] weapons, missile defense weapons, electronic warfare (EW), multiple sensors, and other available weapons/capabilities." (JP 3-01 2012) Active AMD describes such actions as shooting down or diverting incoming missiles and jamming hostile radar or communications.

An example of an active air and missile defense system is the Patriot surface-to-air missile system, which uses an advanced aerial interceptor missile and high performance radar system to detect and shoot down hostile aircraft and tactical ballistic missiles (Patriot 2012). Patriots were first deployed in Operation Desert Storm in 1991 to counter Iraqi Scud missiles. Israel's Iron Dome anti-rocket interceptor system has a similar objective of defending against incoming air threats. According to reports, the system intercepted more than 300 rockets fired by Hamas from Gaza into Israel during the November 2012 conflict, with a success rate of 80 to 90 percent (Kershner 2012). At the time, Israel was also under cyber assault, and Prime Minister Benjamin Netanyahu said that the country needed to develop a cyber defense system similar to Iron Dome (Ackerman and Ramadan 2012).

Another example of an active air defense system is the U.S.'s Operation Noble Eagle (Air Force 2012). Launched minutes after the first aircraft was hijacked the morning of September 11, 2001, the operation has become a major element of

homeland air defense that includes combat air patrols, air cover support for special events, and sorties in response to possible air threats. Although Noble Eagle pilots can potentially shoot down hostile aircraft, so far none have done so. However, they have intercepted and escorted numerous planes to airfields over the years.

In contrast to active defense, passive air and missile defense is defined as: "all measures, other than active AMD, taken to minimize the effectiveness of hostile air and missile threats against friendly forces and assets," noting that "these measures include detection, warning, camouflage, concealment, deception, dispersion, and the use of protective construction. Passive AMD improves survivability by reducing the likelihood of detection and targeting of friendly assets and thereby minimizing the potential effects of adversary reconnaissance, surveillance, and attack." (JP 3-01 2012) Passive AMD includes such actions as concealing aircraft with stealth technology. It covers monitoring the airspace for adversary aircraft and missiles, but not actions that destroy or divert them.

Active and Passive Cyber Defense

We adapt the definitions of active and passive air defense to the cyber domain by replacing the term "air and missile" with "cyber." This gives us the basic definitions: active cyber defense is direct defensive action taken to destroy, nullify, or reduce the effectiveness of cyber threats against friendly forces and assets. Passive cyber defense is all measures, other than active cyber defense, taken to minimize the effectiveness of cyber threats against friendly forces and assets. Put another way, active defenses are direct actions taken against specific threats, while passive defenses focus more on protecting cyber assets from a variety of possible threats.

Using these definitions, we now examine various cyber defenses to see whether they are active or passive. We begin with encryption, which is clearly a passive defense. It is designed to ensure that information is effectively inaccessible to adversaries that intercept encrypted communications or download encrypted files, but takes no action to prevent such interceptions or downloads. Steganography is similarly passive. By hiding the very existence of information within a cover such as a photo, it serves as a form of camouflage in the cyber domain. Other passive defenses include security engineering, configuration monitoring and management, vulnerability assessment and mitigation, application white listing, limiting administrator access, logging, backup and recovery of lost data, and education and training of users. None of these involve direct actions against a hostile threat.

User authentication mechanisms can be active or passive. For example, consider a login mechanism based on usernames and passwords that denies access when either the username or password fails to match a registered user. We consider this passive if no further action is taken against an adversary attempting to gain access by this means. Indeed, the person might try again and again, perhaps eventually succeeding. Now suppose that the mechanism locks the account after three tries. Then it has an active element in that this particular adversary will be unable to gain entry through that account, at least temporarily. However, it does not stop the adversary from trying other accounts or trying to gain access through other means such as a malware attack. Nor does it prevent an attacker who stole an account and password from gaining access to the system.

Now consider DARPA's active authentication program, which seeks to validate users continuously using a wide range of physical and behavioral biometrics such as mouse and typing patterns and how messages and documents are crafted (DARPA 2012). If at any time a user's actions are inconsistent with their normal biometric patterns (called their "cognitive

fingerprint"), access could be terminated. Such a mechanism would be more active than the password mechanism above, as it could keep the adversary from entering and then exploiting any legitimate account on the system. It might even thwart a malware attack, as the malware's behavior would not match that of the account under which it is running.

Consider next a simple firewall access control list (ACL) that blocks all incoming packets to a particular port on the grounds that because the system does not support any services on that port, it would be an open door for attackers. We consider this passive, as it serves more to eliminate a vulnerability than to address a particular threat. However, the ACL would become an element of an active defense if an intrusion prevention system (IPS) detected hostile traffic and then revised the ACL to block the offending traffic. However, an intrusion detection system (IDS) alone is more passive, as it serves primarily as a means of detection and warning.

> active defenses are direct actions taken against specific threats, while passive defenses focus more on protecting cyber assets

Anti-malware (aka anti-virus) tools have much in common with intrusion prevention systems. They detect malicious software, including viruses, worms, and Trojans, and then (optionally) block the code from entering or executing on a protected system. Typically these tools are regularly updated to include signatures for new forms and variants of malware that are detected across the Internet. In this sense, the active defenses are applied globally over the Internet. After new malware is discovered, security vendors create and distribute new signatures to the customers of their anti-malware products.

Intrusion prevention can likewise be performed on a broader scale than a single network or even enterprise. For example, the IP addresses of machines that are spewing hostile packets can be shared widely through "blacklists" and then blocked by Internet service providers. Indeed, victims of massive denial-of-service (DoS) attacks frequently ask upstream service providers to drop packets coming from the originating IP addresses.

Anti-malware and intrusion prevention systems can be integrated to form powerful active defenses. In many respects, the combined defenses would resemble an active air and missile defense system that detects hostile air threats and then takes such actions as shooting them down or jamming their communication, only in cyberspace the defenses are applied to hostile cyber threats such as malicious packets and malware. Rather than targeting incoming ballistic missiles, cyber defenses take their aim at packets that act like "cyber missiles."

Honeypots, which lure or deflect attackers into isolated systems where they can be monitored, are another form of active defense. They are like the decoys used in air defense to deflect missiles away from their intended targets.

In addition to playing a role in network security, active cyber defenses have been used to take down botnets (networks of compromised computers) and counter other cyber threats. The following two examples illustrate.

Coreflood Takedown

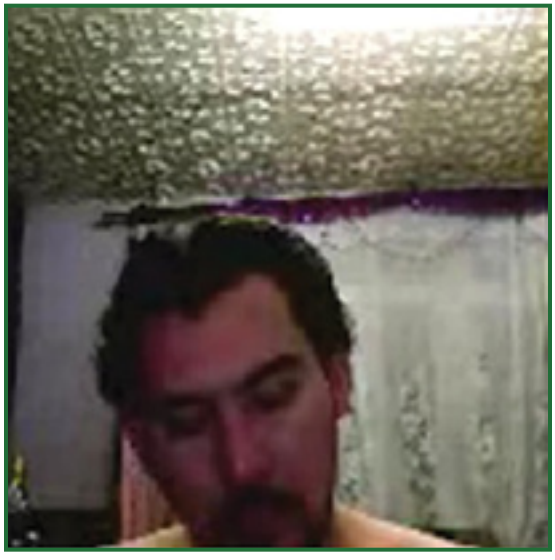
In April 2011, the Federal Bureau of Investigation (FBI), Department of Justice, and the Internet Systems Consortium (ISC) deployed active defenses to take down the Coreflood botnet (Zetter 2011a, 2011b; Higgins 2011). At the time, the botnet comprised over 2 million infected computers, all under the control of a set of command and control (C2) servers. The bot malware installed on the machines was used to harvest usernames and passwords, as well as financial information, in order to steal funds. One C2 server alone held about 190 gigabytes of data stolen from over 400,000 victims.

The active defense included several steps. First, the U.S. District Court of Connecticut issued a temporary restraining order that allowed the non-profit Internet Systems Consortium (ISC) to swap out Coreflood's C2 servers for its own servers. The order also allowed the government to take over domain names used by the botnet. With the infected machines now reaching out to the new C2 servers for instructions, the bots were commanded to "stop." The malware reactivated following a reboot, but each time it contacted a C2 server, it was instructed to stop. The effect was to neutralize, but not eliminate, the malware installed on the compromised machines. To help victims

remove the malware, the FBI provided the IP addresses of infected machines to ISPs so they could notify their customers. In addition, Microsoft issued an update to its Malicious Software Removal Tool, so that victims could get rid of the code.

Using the air defense analogy, the Coreflood takedown can be likened to an active defense against hijacked aircraft, where the hijackers were acting on instructions transmitted from a C2 center. In this situation, the air defense might jam the signals sent from the center and replace them with signals that command the hijackers to land at specified airports. The airports would also be given information to identify the hijacked planes so that when they landed, the hijackers could be removed.

This approach of neutralizing the damaging effects of botnets by commandeering their C2 servers has been used in several other cases. Microsoft, for example, received a court order in November 2012 to continue its control of the C2 servers for two Zeus botnets. Because Zeus had been widely used to raid bank accounts, the operation has no doubt prevented considerable harm (Lemos 2012).



Georgian Outing of Russia-Based Hacker

In October 2012, *Network World* reported that the Georgian government had posted photos of a Russian-based hacker who had waged a persistent, months-long campaign to steal confidential information from Georgian government ministries, parliament, banks, and non-government organization (Kirk 2012). The photos, taken by the hacker's own webcam, came after a lengthy investigation that began in March 2011 when a file on a government computer was flagged by an anti-virus program. After looking into the incident, government officials determined that 300 to 400 computers in key government agencies had been infected with the malware, and that they had acquired it by visiting Georgian news sites that had been infected themselves, in particular, on pages with headlines such as "NATO delegation visit in Georgia" and "U.S.-Georgian agreements and meetings." Once installed, the malware searched for documents using keywords such as USA, Russia, NATO, and CIA, and then transmitted the documents to a drop server where they could be retrieved by the spy.

Georgia's initial response included blocking connections to the drop server and removing the malware from the infected websites and personal computers. However, the spy did not give up and began sending the malware out as a PDF file attachment in a deceptive email allegedly from admin@president.gov.ge.

The Georgian government then let the hacker infect one of their computers on purpose. On that computer, they hid their own spying program in a ZIP archive entitled "Georgian-NATO Agreement." The hacker took the bait, downloaded the archive, and unwittingly launched the government's code. The spyware turned on the hacker's webcam and began sending images back to the government. It also mined the hacker's computers for documents, finding one that contained instructions, in Russian, from the hacker's handler about who to target and how, as well as circumstantial evidence suggesting Russian government involvement.

Again using the air defense analogy, the steps taken to block the exfiltration of files from compromised computers to the drop servers could be likened to jamming the transmission of sensitive data acquired with a stolen reconnaissance plane to the thieves' drop center. The steps taken to bait the hacker into unwittingly stealing and installing spyware might be likened to a command intentionally permitting the theft of a rigged reconnaissance plane with hidden surveillance equipment that sends data it collects about the thieves back to the command.

CHARACTERISTICS AND ETHICAL ISSUES IN ACTIVE CYBER DEFENSE

In this section, we offer a set of distinctions for characterizing the different types of active defense described in the preceding section and discuss some of the ethical issues raised by each.

Scope of Effects

The first set of distinctions pertains to the scope of effects of an active defense. An active defense is said to be internal if the effects are limited to an organization's own internal network. If it affects outside networks, it is said to be external.

Drawing on the air defense analogy, an internal cyber defense is like an air defense system that takes actions against an incoming missile or hostile aircraft after it has entered a country's airspace, while an external cyber defense is like an air defense system that operates in someone else's airspace or attacks the base in a foreign country where the missile is being launched or the hostile aircraft taking off.

We consider defenses that involve sharing threat information with outside parties to be external. An example is the Defense Industrial Base (DIB) Cybersecurity Information Assurance (CS/IA) program operated by the Department of Defense. Under the program, DoD provides DIB companies with unclassified indicators (signatures) of cyber threats. An optional part of the program, called DIB Enhanced Cybersecurity Services (DECS) and run jointly with DHS, allows DoD also to share classified threat information (DoD 2012a).

Most of the effects in the Coreflood takedown were external. In particular, the ISC-operated C2 servers instructed bots in outside networks to stop. In contrast, most of the effects in the Georgian case were internal. Connections to the drop server were blocked on internal networks and internal machines were cleaned of the malware. However, there were also external effects, namely, infection of the hacker's own computer with spyware.

Ethical Issues

In general, most of the ethical issues regarding active defenses concern external active defenses. These will be discussed in the next section when we distinguish cooperative external defenses from non-cooperative ones. However, even internal defenses can raise ethical issues. For example, inside users might complain that their rights to free speech were violated if internal defenses blocked their communications with outside parties. In addition, internal defenses do nothing to mitigate threats across cyberspace. By not even sharing threat information with outsiders, external networks are exposed to continued harm that might be avoided if the defenses were applied to them as well. Arguably, at least in terms of national cyber defense, a better moral choice would be to help mitigate cyber threats more broadly, as DoD has done with its DIB CS/IA and DECS programs. Returning to the air defense analogy, a missile defense system that only shot down missiles headed to military bases would not be as "just" as one that also shot down missiles headed to civilian targets such as cities and malls. On the other hand, it would be unreasonable to expect that missile defense system to protect the air space of other countries, at least absent an agreement to do so.

Degree of Cooperation

The second set of distinctions pertains to the degree of cooperation in an active defense. If all effects against a particular network are performed with the knowledge and consent of the network owner, they are said to be cooperative. Otherwise, they are classified as non-cooperative. For the purpose of discussion here, we assume that network owners are authorized to conduct most defensive operations on their own networks, at least as long as they do not violate any laws or contractual agreements with their customers or users. Thus, the distinction applies mainly to active defenses with external effects.

Using the air defense analogy, a cooperative cyber defense is like an air defense system that shoots down missiles or hostile aircraft in the airspace of an ally that has requested help, while a non-cooperative cyber defense is like an air defense system that shoots them down in the adversary's own airspace.

Anti-viral tools are cooperative defenses. Security vendors distribute new signatures to their customers, but the signatures are only installed with customer permission. Similarly, sharing blacklists of hostile IP addresses is cooperative. In general, any active defense that does nothing more than share threat information is cooperative.

Defenses become non-cooperative when they involve actions taken against external computers without permission of the user or network owner. In the case of Coreflood, the actions taken against the individual bots were non-cooperative. Neither the users of those machines nor the owners of the networks on which they resided agreed to have the bot code stopped. On the other hand, neither had they agreed to the initial malware infection and subsequent theft of their data. Arguably, any user would prefer that the malware be stopped rather than be allowed to continue its harmful actions. Further, even though the action was non-cooperative, it was deployed under legal authorities, enabled in part by the temporary restraining order. Moreover, the actual elimination of the malware from the infected machines was to be a cooperative action involving the machine owners.

Non-cooperative defenses include what is sometimes called "attack back," "hack back," or "counterstrike" where the defense uses hacking or exploit tools directly against the source of an attack or gets the attacker to unwittingly install software, say by planting it in a decoy file on a computer the attacker has compromised. The goal might be to collect information about the source of the attack, block attack packets, or neutralize the source. Non-cooperative defenses also include court-ordered seizures of computers.

Although the Coreflood takedown did not include any sort of hack back, the Georgian case did. In particular, the actions taken to plant spyware on the hacker's computer constituted a non-cooperative counterstrike. However, one could argue that the hacker would never have acquired the spyware had he not knowingly and willfully first infected the computer hosting it and second downloaded the ZIP archive containing it. Thus, he was at least complicit in his own infection and ultimate outing.

Ethical Issues

As a rule, non-cooperative defenses, particularly those involving some sort of hack back, raise more ethical and legal issues than cooperative ones. In part, this is because most cyber attacks are launched through machines that themselves have been attacked, making it hard to know whether the immediate source of an attack is itself a victim rather than the actual source of malice. They may be hacked servers or bots on a botnet. Thus, any actions taken against the computers could harm parties who are not directly responsible for the attacks. In addition, cyber attacks in general violate computer crime statutes, at least when conducted by private sector entities. While the argument can be made that some hack backs would be permissible under the law, not everyone agrees, and the topic has been hotly debated (Denning 2008, Steptoe 2012, Messmer 2012). However, government entities, in particular the military and law enforcement and intelligence agencies, have or can acquire the authorities needed to perform actions that might be characterized as hacking under certain prescribed conditions.

If we assume that non-cooperative defenses are conducted by or jointly with government entities with the necessary legal authorities, then the primary concern is that innocent parties may be harmed. Then we can draw on the long tradition of just war theory to determine the conditions under which active cyber defenses that pose risks to noncombatants can be ethically justified.

Most just war theorists hold that noncombatant immunity is a key linchpin to all our moral thinking in war (Walzer 1977, Nagel 1972, Rodin 2003, Orend 2006). As such, noncombatants are never to be intentionally targeted for harm as any part of a justified military action. Traditional just war theory does hold, however, that some actions that will foreseeably but unintentionally harm noncombatants may be permissible, so

long as that harm is truly unintentional, is proportionate to the good goal achieved by the act, and is not the means itself to achieve the good goal. Grouped together, these principles are known as the doctrine of double effect. The doctrine has come under heavy scholarly debate, with many critics doubting that its principles can hold true for all cases (Davis 1984, Kamm 2004, McIntyre 2001, Steinhoff 2007), while others have argued that some revised or narrowed version of the doctrine can still be defended and applied to war (McMahan 1994, Quinn 1989, Nelkin and Rickless 2012). We cannot here engage this larger debate, but assume that at least some narrow version of the doctrine of double effect is true and, as such, it is critical for our moral conclusions regarding harm to noncombatants from active cyber defense.

Whether noncombatants' property can be targeted is another matter. Generally, noncombatant property is similarly considered immune from direct and intentional harm since harming one's property harms that person. However, as with physical harm, unintended harm of noncombatant property can be permissible in some instances. Moreover, traditional just war theory and the laws of armed conflict can allow for some level of intentional harm to civilian property if it is necessary to block a particularly severe enemy military action and the civilians in question are later compensated. That is, generally, the ethical restrictions on harm to civilian property are far less strict than for physical harm to civilian persons. This is true for unintentional harms of both kinds, and can even allow for some intentional harm to property when necessary, the stakes are high enough, and recompense can be made.

In the case of active air defense, systems like Iron Dome are not without risk to civilians. If someone happens to be under an incoming rocket's flight path at the time it is hit, they could be harmed from fallout from the explosion. However, Israel has limited their counterstrikes primarily to rockets aimed at densely populated urban areas. In that situation, any fallout is likely to be substantially less harmful than the effects produced by the rockets themselves if allowed to strike. We argue that such a risk imposition can be morally warranted. Note, however, that if Iron Dome created large amounts of dangerous and lethal fallout disproportionate to the lives saved, then its use would not be permissible.

In general, if an air defense system distributes some small level of risk of harm to civilians under an incoming missile's flight path in order to protect a much larger number of civilians from much greater harm, then the conditions are present for such defense to be morally permissible. This is precisely what we find in the case of real-world air defense systems such as Iron Dome. Further, it is irrelevant whether the risk of harm is imposed on noncombatants from one's own state or another state. The reason is that what matters are the moral rights of *all* noncombatants, including, of course, noncombatants on either side of a given conflict. The point is to minimize collateral harm to all noncombatants.

The same principles should apply to active cyber defense; that is, it should be morally permissible for a state to take an action against a cyber threat if the unjust harm prevented exceeds and is proportionate to any foreseen harm imposed on noncombatants. Indeed, in the cyber domain it will often be easy to meet this demand because it is often possible to shoot down the cyber missiles without causing any fallout whatsoever. Instead, packets are simply deleted or diverted to a log file. Nobody is harmed.

In some cases, however, an active defense could have a negative impact on innocent parties. To illustrate, suppose that an action to shut down the source of an attack has the effect of shutting down an innocent person's computer that had been compromised and used to facilitate the attack. In this case, the action might still be morally permissible. There are two reasons. First, the harm induced might be temporary in

nature, affecting the computer, but only for a short time until the attack is contained. Second, the harm itself might be relatively minor, affecting only noncombatants' property, not their persons. It is possible that such effects could further impede other rights of the noncombatants, such as their ability to communicate or engage in activity vital to their livelihoods. But all of these further harms would be temporary in nature and could even be compensated for, if appropriate, after the fact. This is not to disregard the rights of noncombatants and their property and its use for the furtherance of other rights in our moral calculus, but is rather a simple recognition that different kinds and severities of harm result in different moral permissions and restrictions.

The fact that the harm itself is likely to be non-physical is quite significant in our moral reasoning in favor of active cyber defense. If it is permissible in some cases to impose the risk of *physical* harm on noncombatants as part of a necessary and proportionate defensive action against an incoming missile (as we argued above that it could be in the air defense case), then surely there will be cases where it can be permissible to impose the risk of temporary harm to the *property* of noncombatants in order to defend against an unjust cyber attack. The point here with active cyber defense is that the *kind* of harms that would be potentially imposed on noncombatants, in general, are the kinds of reduced harms that should make such defensive actions permissible.

A caveat, however, is in order. Computers today are used for life-critical functions, for example, to control life support systems in hospitals and operate critical infrastructure such as power grids. In a worst-case, an active cyber defense that affects such a system might lead to death or significant suffering. These risks need to be taken into account when weighing the ethics of any non-cooperative action that could affect noncombatants. In general, defensive actions that do not disrupt legitimate functions are morally preferable over those that do. If the scope of possible effects cannot be reasonable estimated or foreseen, then the action may not be permissible.

In the case of Coreflood, the takedown affected many noncombatant computers. However, the effect was simply to stop the bot code from running. No other functions were affected, and the infected computer continued to operate normally. Thus, there was virtually no risk of causing any harm whatsoever, let alone serious harm. In the Georgian case, the only harm was to the attacker's own computer—and he brought this on himself by downloading the bait files, thus making himself liable to intentional defensive harm.

Although the discussion here has focused on non-cooperative defenses, it is worth noting that while cooperative defenses generally raise fewer issues, they are not beyond reproach. For example, suppose that a consortium of network owners agrees to block traffic from an IP address that is the source of legitimate traffic as well as the hostile traffic they wish to stop. Depending on circumstances, a better moral choice might be to block only the hostile traffic or work with the owner of the offending IP address to take remedial action.

While cooperative defenses generally raise fewer issues, they are not beyond reproach.

Types of Effects

The third set of distinctions pertains to the effects produced. An active defense is called sharing if the effects are to distribute threat information such as hostile IP addresses or domain names, or signatures for malicious packets or software, to other parties. Sharing took place in the Coreflood takedown when the FBI provided the IP addresses of compromised machines within the United States to their U.S. ISPs and to foreign law enforcement agencies when the machines were located outside the U.S.. Another example of sharing is DoD's DIB program, described earlier.

An active defense is called collecting if it takes actions to acquire more information about the threat, for example, by activating or deploying additional sensors or by serving a court order or subpoena against the source or an ISP likely to have relevant information. In the Coreflood takedown, the replaced C2 servers were set up to collect the IP addresses of the bots so that eventually their owners could be notified. The servers did not, however, acquire the contents of victim computers. In the Georgian case, spyware was used to activate a webcam and collect information from the attacker's computer.

An active defense is called blocking if the effects are to deny activity deemed hostile, for example, traffic from a particular IP address or execution of a particular program. The Coreflood takedown had the effect of breaking the communication channel from the persons who had been operating the botnet to the C2 servers controlling it. As a result, they could no longer send commands to the bots or download stolen data from the servers. In the Georgian case, connections to the drop servers were blocked in order to prevent further exfiltration of sensitive data.

Finally, an active defense is pre-emptive if the effects are to neutralize or eliminate a source used in the attacks, for example, by seizing the computer of a person initiating attacks or by taking down the command and control servers for a botnet. In the Coreflood takedown, the hostile C2 servers were put out of commission and the bots neutralized. With further action on the part of victims, the malware could also be removed.

Using the air defense analogy, the cyber defense of sharing is like a missile defense system that reports new missile threats to allies so that they can shoot them down. The cyber defense of collecting is like a missile defense system that activates additional radars or other sensors in response to an increased threat level, or that sends out sorties to investigate suspicious aircraft. The cyber defense of blocking is akin to a missile defense system that shoots down incoming missiles or jams their radars and seekers. Finally, the cyber defense of pre-emption is like launching an offensive strike against the air or ground platform launching the missiles.

Some authors regard retaliation or retribution as a form of active defense. However, we consider these operations to be offensive in nature, as they serve primarily to harm the source of a past attack rather than mitigate, stop, or pre-empt a current one.

Ethical Issues

All four types of cyber operations raise ethical issues. The act of sharing raises issues of privacy and security, particularly if any sensitive information is shared. The act of collecting also raises issues about privacy and security, but in this case relating to the new information acquired rather than the dissemination of existing information. The act of blocking raises issues relating to free speech and over-blocking. In a worst case, traffic might be blocked that is important for the operation of a life-support system or critical infrastructure such as power generation and distribution. Likewise, the act of pre-emption raises ethical issues relating to disabling software or systems. Again, a worst-case scenario could cause serious harm, for example, by shutting down a life-support system. These possible harms would need to be considered in the application of any non-cooperative cyber defense, as discussed in the previous section, and argue for defenses that limit their effects, say, by disabling only traffic and software involved in an attack rather than shutting down all traffic and complete systems.

In the Coreflood takedown, it is important to note that the government did not attempt to remove the bot code from infected machines. They only neutralized it by issuing the stop command. Part of the reason for not removing the code was a concern for unanticipated side effects that might damage an infected computer.

Because active cyber defense is a form of defense that should not be misconstrued as offense, it is worth explaining why the distinction between offensive retaliation versus legitimate defensive action is so crucial in the ethical dimensions of killing and war. Defensive harm has the lowest ethical barrier to overcome from amongst all possible justifiable harms. That is, if one is being wrongly attacked, then the moral restrictions against using force of some kind in order to block that wrongful attack are (relatively) few. This is because all people have a

right not to be harmed unjustly. If one is attempting to harm someone unjustly, then she has made herself morally liable to suffer defensive harm as part of an act taken to thwart her attempted unjust harm. The person being wrongly attacked may permissibly harm his attacker in an effort to block or thwart the attack against him, so long as the defensive harm meets two criteria. First, it must be necessary to inflict the defensive harm to block the unjust attack. If the defensive harm in question does nothing to block the liable party's unjust attack, then it is retributive punishment, or something else, but not properly an act of defense. Second, the defensive harm must be proportionate to the unjust harm to be blocked. If a foreign plane was found conducting reconnaissance over a state's territory without permission during peacetime, then the foreign state may have made itself liable to some form of defensive action such as being escorted to an airfield. However, it would be disproportionate and wrongful to shoot the plane down or, even worse, shoot down commercial planes flying under the foreign state's flag. In general, there must be some reasonable correlation and proper "fit" between the extent of defensive response and the degree of liability of the party defended against (McMahan 2005, Quong 2012). In the case of an active cyber defense, if the act is truly a defensive effort to block an unjust attack, then so long as it is necessary and proportionate, it will usually be ethically permissible. In the Georgian case, the government responded to the cyber espionage operation against it with its own espionage operation against the hacker. It did not destroy software and data on the hacker's computer.

Degree of Automation

The final set of distinctions pertains to the degree of human involvement. An active defense is said to be automatic if no human intervention is required and manual if key steps require the affirmative action of humans.

Most anti-malware and intrusion prevention systems have both manual and automated components. Humans determine what goes into the signature database, and they install and configure the security software. However, the processes of signature distribution, malicious code and packet detection, and initial response are automated.

In the Coreflood takedown, the execution of the stop commands was fully automated through the C2 servers. However, humans played an important role in planning and decision making, analyzing the botnet code and the effects of issuing a stop command, acquisition of the restraining order, and swapping out of the C2 servers. Thus, the entire operation had both manual and automatic aspects. In the Georgian case, much of the investigation involved manual work, including analyzing the code, determining what the hacker was looking for, and setting up the bait with the spyware. But the key element in the outing, namely the operation of the spyware, was automated. Once the hacker downloaded the ZIP archive, it did the rest.

Applying the air defense analogy once again, an automatic cyber defense is like a missile defense system that automatically shoots down anything meeting the preset criteria for being a hostile aircraft or incoming missile, whereas a manual cyber defense is more like Operation Noble Eagle where humans play a critical role, both in recognizing and responding to suspicious activity in U.S. airspace.

Ethical Issues

In general, manual actions give humans a greater opportunity to contextualize their ethical decisions. Rather than configuring a system to respond always in a certain way, humans can take into account the source or likely source of a perceived threat, its nature, and the likely consequences of taking certain actions against it. This is vital to Nobel Eagle, where most incidents

turn out to be non-hostile and lives are at stake. On the other hand, manual actions take longer to execute than automated ones, potentially allowing greater damage to be incurred before the threat is mitigated. In the cyber domain, where actions can take place in an instant, automated defenses become critical. That is, the speed of some actions in the cyber domain are such that a cyber defense must be automated in order to have any effect at all against the attack. It is perhaps for this reason that some cyber actions have been given an exemption from the recent "man in the loop" legal requirements for automated weapon systems put out by the DoD (DoD 2012b, Gallagher 2012). If a hostile actor has launched an attack to cause a power generator to explode, then an automated response that successfully blocks the attack without causing unnecessary harm is morally superior to a manual one that comes too late.

However, this does not mean that all cyber defenses should be automated. To be clear: we are not arguing that *all* cyber actions should be exempt from the "man in the loop" requirement. The nature of a defense and its potential effects must be weighed in any decision to automate. The potential severity of foreseeable harms should govern whether it should be automated. It is true that the cyber case is unique in that the speed of many cyber attacks necessitates that many defenses be automated in order to be effective in any way. But if the effects of a given defense are such that their automation would lead to too great a risk of impermissible harm, then they should not be automated, even if this entirely nullifies their efficacy. Thankfully, given the reasons discussed above regarding the kinds of predictable effects that most forms of active cyber defense would result in, we find that in many cases their automation could be permissible.

Conclusions

Using analogies from air defense, we have shown that active cyber defense is a rich concept that, when properly understood and executed, is neither offensive nor necessarily harmful and dangerous. Rather, it can be executed in accordance with the well-established ethical principles that govern all forms of defense, namely principles relating to harm, necessity, and proportionality. In many cases, such as with most botnet takedowns, active defenses mitigate substantial harm while imposing little or none of their own.

While active defenses can be morally justified in many cases, we do not mean to imply that they always are. All plausible effects must be considered to determine what, if any, harms can follow. If harms cannot be estimated or are unnecessary or disproportionate to benefits gained, an active defense cannot be morally justified.

In considering active defenses, we have assumed that they would be executed under appropriate legal authorities. In particular, they would be conducted by authorized government entities or by private companies operating under judicial orders or otherwise within the law. We leave open the question of how far companies can go in areas where the law is unclear or untested. While such active defenses as sharing attack signatures and hostile IP addresses and domain names have raised few legal questions, an active defense that deleted code or data on the attacker's machine would raise more. No doubt, this area will likely continue to inspire lively discussion and debate. ❊

REFERENCES

Ackerman, G. and Ramadan, S. A. (2012) 'Israel Wages Cyber War With Hamas as Civilians Take Up Computers,' *Bloomberg*, November 19. http://www.bloomberg.com/news/2012-11-19/israel-wages-cyber-war-with-hamas-as-civilians-take-up-computers.html (accessed November 26, 2012).

Air Force. (2012) Operation Noble Eagle, Air Force Historical Studies Office, Posted September 6. http://www.afhso.af.mil/topics/factsheets/factsheet.asp?id=18593 (accessed November 6, 2012).

DARPA. (2012) 'Active Authentication,' DARPA Information Innovation Office, http://www.darpa.mil/Our_Work/I2O/Programs/Active_Authentication.aspx (accessed November 6, 2012).

Davis, N. (1984) 'The Doctrine of Double Effect: Problems of Interpretation,' *Pacific Philosophical Quarterly* 65: 107–123.

Denning, D. E. (2008) 'The Ethics of Cyber Conflict,' Chapter 17 in *The Handbook of Information and Computer Ethics* (K. E. Himma and H. T. Tavani eds.), Wiley, pp. 407–428.

Dittrich, D. and Himma, K. E. (2005) 'Active Response to Computer Intrusions,' *The Handbook of Information Security* (Bidgoli, H. ed.), John Wiley & Sons.

DoD. (2012a) Fact Sheet: Defense Industrial Base (DIB) Cybersecurity Activities, May 11, 2012. http://www.defense.gov/news/d20120511dib.pdf (accessed December 5, 2012).

DoD. (2012b) Directive Number 3000.09, 'Autonomy in Weapon Systems,' November 21, 2012. www.dtic.mil/whs/directives/corres/pdf/300009p.pdf (accessed December 6, 2012).

Gallagher, S. (2012) 'U.S. cyber-weapons exempt from "human judgment" requirement' *arstechnica,* November 29, 2012. http://arstechnica.com/tech-policy/2012/11/us-cyber-weapons-exempt-from-human-judgment-requirement/ (accessed December 5, 2012).

Higgins, K. J. (2011) 'Coreflood Botnet An Attractive Target For Takedown For Many Reasons,' *Dark Reading*, April 14. http://www.darkreading.com/database-security/167901020/security/client-security/229401635/coreflood-botnet-an-attractive-target-for-takedown-for-many-reasons.html (accessed November 27, 2011).

JP 3-01 (2012) 'Countering Air and Missile Threats,' Joint Publication 3-01, March 23.

Kamm, F. (2004) 'Failures of Just War Theory: Terror, Harm, and Justice,' *Ethics* 114: 650–92.

Kershner, I. (2012) 'Israeli Iron Dome Stops a Rocket With a Rocket,' *The New York Times*, November 18. http://www.nytimes.com/2012/11/19/world/middleeast/israeli-iron-dome-stops-a-rocket-with-a-rocket.html?_r=0 (accessed November 19, 2012).

Kirk, J. (2012) 'Irked By Cyberspying, Georgia Outs Russia-Based Hacker—With Photos,' *Network World*, October 30. http://www.networkworld.com/news/2012/103012-irked-by-cyberspying-georgia-outs-263790.html (accessed November 27, 2012).

Lemos, R. (2012) 'Microsoft Can Retain Control of Zeus Botnet Under Federal Court Order,' eWeek, December 1. http://www.eweek.com/security/microsoft-can-retain-control-of-zeus-botnet-under-federal-court-order/ (accessed December 7, 2012).

McIntyre, A. 'Doing Away with Double Effect,' *Ethics* 111: 219–55.

McMahan, J. (1994) 'Revising the Doctrine of Double Effect,' The Journal of Applied Philosophy 11(2): 1993-221.

McMahan, J. (2005) 'The Basis of Moral Liability to Defensive Harm,' *Philosophical Issues* 15: 386–405.

Messmer, E. (2012) 'Hitting Back at Cyberattackers: Experts Discuss Pros and Cons,' *Network World*, November 1. http://www.networkworld.com/news/2012/110112-cyberattackers-263885.html (accessed November 29, 2012).

Nagel, T. (1972) 'War and Massacre,' *Philosophy and Public Affairs*, 1(2): 123–144.

Nelkin, D. and Rickless, S. (2012) 'Three Cheers for Double Effect,' *Philosophy and Phenomenological Research,* http://onlinelibrary.wiley.com/doi/10.1111/phpr.12002/full (accessed December 5, 2012).

Orend, B. (2006) *The Morality of War*, Peterborough, ON: Broadview Press.

Patriot. 'MIM-104 Patriot,' Wikipedia. http://en.wikipedia.org/wiki/MIM-104_Patriot (accessed November 6, 2012).

Quinn, W. S. (1989) 'Actions, Intentions, and Consequences: The Doctrine of Double Effect,' *Philosophy and Public Affairs* 18:334–51.

Quong, J. (2012) 'Liability to Defensive Harm,' *Philosophy & Public Affairs*, 40(1): 45–77.

Rodin, D. (2003) *War and Self-Defence*, New York: Oxford University Press.

Steinhoff, U. (2007) *On the Ethics of War and Terrorism*, Oxford: Oxford University Press.

Steptoe. (2012) 'The Hackback Debate,' Steptoe Cyberblog, November 2. http://www.steptoecyberblog.com/2012/11/02/the-hackback-debate/ (accessed November 29).

Walzer, M. (1977) *Just and Unjust Wars*. New York: Basic Books

Zetter, K. (2011a) 'With Court Order, FBI Hijacks Coreflood Botnet, Sends Kill Signal," *Wired*, April 13. http://www.wired.com/threatlevel/2011/04/coreflood/ (accessed November 27, 2012).

Zetter, K. (2011b) 'FBI vs. Coreflood Botnet: Round 1 Goes to the Feds,' *Wired*, April 26. http://www.wired.com/threatlevel/2011/04/coreflood_results/ (accessed November 27, 2012).

NOTES

1 Approved for public release; distribution is unlimited. The views expressed in this document are those of the authors and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

The Strategy of Economic Warfare: A Historical Case Study and Possible Analogy to Contemporary Cyber Warfare

Nicholas A. Lambert[1]

The international economy today bears an uncanny resemblance to that of a century ago. The stability of the modern international economy rests upon the free movement around the globe of goods, money and above all information (broadly defined). Within this globalized trading system, there exist numerous parallels to the 1914 setting: ever-increasing velocity of transactions and pace of economic activity flowing through financial systems; dependency upon accurate and instantaneous information; vulnerability of merchants to a collapse of the insurance and reinsurance industry; and the buying and selling of ever more sophisticated and intangible financial instruments. In all these respects, the world of 2014 resembles that of 1914.

The historical analogy between British economic warfare in the First World War and the prospect of cyberwar in the 21st century is attractive for several reasons. Most obviously, the British sought to wage the war by controlling and carefully disrupting the information systems at the heart of the global economic system, just as cyber warfare today would seek to target information systems at the heart of the global economy. More generally, the analogy's appeal stems from a widespread perception of similarities in the overarching geo-political-economic landscape of the two eras. A century ago, like today, policy planners quite justifiably perceived themselves to be confronted with a very new, almost alien, strategic environment. The development of instantaneous global communications thanks to the spread of the cable (and wireless) telegraph network changed the structure of the world economy in ways that presented multiple challenges and opportunities to those nations that were economically, politically and militarily powerful. No less importantly, yet sometimes overlooked, the effects of this transformation in the world economic system extended beyond states and their militaries and included the lives of individual users. Globalization modified the day-to-day behavior of the average business and consumer, which in turn helped transform the strategic environment.

The analogy also serves as a reminder of how serious the stakes can be when warfare—cyber or otherwise—disrupts the global trading system and thereby causes significant economic dislocation on a scale capable of producing major social upheaval. Throughout the First World War, the stakes were no less than the prospect of social revolution. Within the British government, this fear was acute at the outset, and although within months concern subsided, it never went away, quickly returning to the forefront at times of domestic crisis (such is in early 1915). In 1917, after the Russian revolution, it gained new impetus. British corporations apparently took the prospect of social upheaval so seriously that they gradually came to accept more state regulation in the name of national security as an alternative to the prospect of total confiscation of property in the name of a radicalized society. Their accommodation raised fundamental questions about the relationship between state and society which remain largely unanswered.

Before plunging into the historical analogy between cyber warfare and Britain's plan for and conduct of economic warfare, I should like to make a few disclaimers by way of framing the discussion.

First, until we have developed a 21st-century doctrine of cyberwar, we can only have a rudimentary understanding of the contemporary end of our analogy between past and present.

Second, to the extent that it works, the analogy serves as more of a cautionary tale than a model for emulation. The British government's experiment was beset with unintended consequences and unfulfilled expectations, and fell far short of the goals that planners had established.

Third, economic warfare represented the cornerstone of British grand strategy in 1914; it was not a subsidiary, disruptive or auxiliary strategy. To work, the analogy presupposes that, as in 1914, cyber warfare occurs as the central strategy in peer-to-peer warfare, when in fact it could occur in other ways as well.

The point here is not to present the economic warfare analogy as rigid or narrow, but rather to emphasize that the discussion of cyber warfare below is broad and worst-case, covering only an attack that threatens national social and economic survival. Complete national collapse—not a mere diversion of the German military or government—is what the British strategy of economic warfare aimed at very explicitly in 1914, albeit for just a couple of weeks.

Fourth, economic warfare as the British conceived of it was more than a matter of targeting a specific industry or element of national critical infrastructure. We are not talking about bombing ball-bearing plants or oil refineries (with precision or otherwise), as in World War Two. While waging economic warfare could produce knock-on systemic consequences—creating bottlenecks and chokepoints in terms of stocks and flows—the aim was not to throttle military industry. The aim was far higher, to score a knock-down blow that obviated the need for less intense but more prolonged types of war. Put another way, economic warfare transcended specific systems—it was not intended to be systems specific, but society specific. Indeed, Britain's plan for economic warfare may well have been the first in history to seek victory by deliberately targeting the enemy society rather than the enemy state.

"cyberspace" includes beliefs and psychology, not just electronic systems

Finally, I do not mean to suggest that there exist direct parallels with the British (or German, or American) experience of World War One down to every last detail. Nor do I mean to suggest that the technologies are similar, for in fact they are quite different. Rather, the points that I want to emphasize—the questions I am going to raise—pertain to the economic, political, and legal implications of waging warfare within a globalized trading system, and the difficulties and dangers of trying to weaponize any of the underpinning infrastructure. As the British discovered in 1914, it was easier said than done.

There are 4 basic parts to the story:

First, *why* would one choose to "weaponize" the international trading system in the first place? We must understand how Britain came up with the strategy of economic warfare and why some British planners thought it would work and others thought the concept too dangerous. This question pertains to the strategic environment created by globalization, and must consume a great part of the analogy.

Second, *how* did British strategists go about implementing their strategy? Clearly the technologies will be different from those of today, but if we bear in mind that "cyberspace" includes beliefs and psychology, and not just electronic systems, then we can begin to see how the British conceptualized their offensive and think about some functional requirements or opportunities.

Third, we must look at the consequences, however unexpected or underestimated. What happened? How long did the strategy last? How did it evolve?

Which leads us to our final basic question: Can one prepare to defend as well as to attack? What are some inherent risks and opportunities for defense against economic warfare, and how do they relate to an offensive strategy? How does a state prepare to "endure" economic warfare, as opposed to preparing to "wage" economic warfare?

THE GLOBALIZED ECONOMIC SYSTEM

At the beginning of the twentieth century, there existed widespread recognition that the world was experiencing an epochal transformation. Historians have written much about the changes of the nineteenth century: of agricultural and industrial revolutions, of transportation and communications revolutions, of accelerating urbanization, of the spread of market capitalism across the globe and growth in world trade, and of the political and societal revolutions—that is to say, fundamental changes in the relationship between society and state. But relatively few have considered the consequent transformation in the structure of the world economic system. Fewer still have followed these changes down to the nuts-and-bolts level, to understand their impact upon the day-to-day 'economic' lives of individual citizens or the cumulative consequences of those individual changes for social stability. And even fewer still—arguably none—have considered the

strategic implications of globalization as such, since no such category of analysis existed.

During the last quarter of the nineteenth century, technological innovation led to a series of sharp drops in the cost of transportation by land (railways) and by sea (steamships). Combined with other innovations in financial services (credit financing) and the creation of a communications network that permitted instantaneous communication between almost any two points on the globe (cable later supplemented by wireless), between 1870 and 1913, the volume of world trade doubled and redoubled. These three industries—transportation, financial services and communications—were the central pillars supporting the global trading system: they represented the critical infrastructure for its operation. During this same period, all nations, especially the industrialized European powers, saw a steady rise in the ratio of foreign trade to national economic output. Transnational linkages and interdependencies multiplied and intensified to the point where one could legitimately talk about a global trading system, functioning on a worldwide basis. Commercial supply chains began to stretch around the world, and national economies became intertwined.

These macro-economic changes were accompanied by astonishingly rapid changes in micro-economic behavior, in other words the day-to-day methods by which international (overseas) commerce was conducted. The net effect of these changes was the creation of a new commercial environment. This may be attributed to a number of factors but unquestionably the most important of these, the one that more than any other defined the essential character of the new economic world system, was the creation during the 1870s of the global cable network. The speed with which commerce was negotiated and transacted, the velocity of transactions, accelerated beyond all experience. Railways and steamships enabled the delivery of goods faster and cheaper than ever before; this much is obvious. But the advent of a cheap and reliable means of real-time communication between buyers and vendors—instead of letters sent by sailing ships—was fundamental.

Even more remarkable and far reaching, however, was the extent to which the structure of commerce became reoriented to the cable. Within a remarkably short period, just twenty-five years or so, most international (and much domestic) commerce became reliant upon access to cable communications: to allow buyers and vendors to find each other in the first place; to negotiate contracts; to arrange credit financing (a bill of exchange drawn on a London bank), insurance and shipping; to schedule payment and final delivery. It is tempting here to draw a parallel to the equally rapid spread and adoption of the Internet, and in some respects one may legitimately do so, but at the same time it is essential always to consider "modern" developments in light of what existed before—to measure relative not absolute change. Relative change in the earlier era was far greater than today.

By the turn of the twentieth century, instantaneous communications had transformed the day-to-day conduct of international trade. It was increasingly commonplace for merchantmen to load with a bulk commodity (say wheat), clear port in the Americas and proceed east without an ultimate destination. Cargoes in transit could—and frequently did—change ownership during the voyage. Only after crossing the Atlantic, a voyage that took approximately a fortnight, and touching at Falmouth (UK) to refuel might the master receive instructions on where to discharge the cargo (be it London, Rotterdam, or Hamburg). With the advent of wireless telegraphy during the first decade of the twentieth century, merchantmen still at sea could be diverted in order to maximize profit.

While these developments increased commercial flexibility and efficiency and thus lowered economic costs, at the same time they created profound problems for traditional methods of waging maritime war. For instance, there were serious legal implications. The entire structure of international maritime law pertaining to war at sea, and specifically the law of blockade, was based on older trading practices. In particular, the law assumed that a belligerent—its navy and its prize courts—could know the ultimate destination of a merchant vessel and ownership of its cargo. In the age of real-time communications, however, and subsequent changes in business norms, it was frequently impossible to obtain that knowledge, because even the master of the vessel in question often did not know his ultimate destination until late in the journeys and because ownership of cargoes often changed hands while vessels were at sea. Even at the point of unloading a cargo might not necessarily have a clear owner. In fact, no means or mechanism, municipal or international, existed anywhere to verify the ownership or destination of merchant ship cargoes. The implications were staggering. In time of war, the immutable rights of neutrals under international law to maintain their legitimate trade had become fundamentally irreconcilable with the equally immutable rights of belligerents to prevent illegitimate contraband from reaching their enemies.

Without near-perfect information, there would be no "global price"

Such changes in the day-to-day conduct of international trade also had enormous strategic implications, which require some grasp of classical economic theory to understand. In classical economics, conditions of near-perfect information in any market must be a pre-requisite for near-perfect competition, which encourages prices to converge. After about 1870, the global market began to mimic economic theory. In other words, only after 1870 did real-time global communications permit near-perfect "global" information, and only after 1880 did low-cost shipping make it possible for merchants to exploit that information both quickly and cheaply. As a result of the development of a global information network, therefore, the prices of staple commodities between different regional markets around the globe became increasingly integrated, reflected in the convergence of prices. In 1870, for instance, the price of wheat in Liverpool exceeded Chicago prices by 57.6%, in 1895 by 17.8%, and in 1913 by just 15.6%. Similar trends may be observed between other commodities and between other regional markets, such as Liverpool and Odessa. (Incidentally, these trends are equally evident within national markets: in 1870, the wheat price spread between New York City and Iowa was 69%; by 1910 it had fallen to just 19%.)

Of course, staples like wheat had been traded across oceans for centuries if not millennia; but the phenomenon of price convergence described above was something radically new; the necessary technological and information apparatus for allowing staples to move fluidly between regional markets, and in such volume, simply did not exist before.

Economic theory also calls for economic behavior to adjust to the convergence of prices (as indeed it calls for behavior to adjust to any change in prices). Again, reality mimicked theory, as economic actors around the globe, ranging from nation-states to individual households, began to plan their budgets and make investments (for instance to expand production) and purchases based on certain expectations derived from knowledge of a "global" price. One of the most significant indicators of these changes was the proliferation of "just-in-time" ordering—that is, not laying up reserves in advance ("buffer stocks") and thereby incurring storage and other costs, due to uncertainty about prices and the inability to communicate in real time with sellers, but waiting to order goods until "just in time" on the expectation that their prices would remain stable and that it would be possible to communicate with sellers at the last minute. Already by the turn of the twentieth century, stocks of wheat in Britain were being measured in terms of weeks rather than months of supply.

The fact that economic activity, at all levels, had grown dependent upon access to near-perfect information created enormous new strategic opportunities and vulnerabilities. Of these, the most important was that it became possible for economic pressure, traditionally applied by navies rather than armies, to target whole societies and collapse them rapidly rather than merely to target state revenues and erode them slowly. The salient change was not that the chief source of state revenue

in previous centuries vulnerable to naval pressure had been international trade (the other chief source was land, which was not vulnerable to naval pressure), and that the sources of state revenues were becoming more diversified (although they were). Rather, it was that the economic well-being of whole societies, not merely governments, depended upon a highly optimized economic system, itself dependent upon access to the infrastructure of global trade, reliant upon access to real-time communications.

The essential point here is that across the spectrum of economic activity—national, business and individual—economic behavior had grown dependent upon the availability and flow of "perfect information". Particularly in industrialized nation-states with large urban populations, social and political stability had come to require high levels of economic prosperity, made possible by a steady supply (flow) of goods and staples through the global trading system. At the most basic bio-physical level also, urban populations needed a steady supply of food as they had no reserves of their own; cities contained no stockpiles beyond what was on the shelves, at most enough to last for a few weeks. Producers, meanwhile, depended on selling their produce in a constant stream of commerce. If it piled up on the wharves, they would be in deep trouble, as would the banks that had extended loans to them. To be clear, we are talking about two related but distinct types of fragility: both that of the economic system and that of politically-aware industrial societies. Of course, given time, businessmen could return to the old ways or otherwise adapt—but time was quite literally of the essence. The key question here, which pre-1914 statesmen across Europe pondered and worried about, was: how much time was needed for businesses and economies to adapt? And, of course, what would happen in the interim?

Of course, information was not the only part of the global economic system on which whole societies increasingly depended. Particularly in industrialized nation-states with large urban populations, social and political stability had come to require high levels of economic prosperity, made possible by a steady supply (flow) of goods and staples through the global trading system. At the most basic bio-physical level also, urban populations needed a steady supply of food as they had no reserves of their own; cities contained no stockpiles beyond what was on the shelves, at most enough to last for a few weeks. Producers meanwhile depended on selling their produce in a constant stream of commerce. If it piled up on the wharves, they would be in deep trouble, as would the banks that had extended loans to them. These conditions made possible a peculiar scenario: if supply were interrupted, then shortages could occur even in a year with a bumper crop; consequently, farmers could suffer a plunge in commodity prices while consumers faced a surge in retail prices. In short, we are talking about two related but distinct types of fragility: both that of the economic system *and* that of politically-aware industrial societies.

This new economic-commercial environment bore a resemblance to the present wave of globalization in that it constituted a dramatic acceleration and intensification of the flow of information throughout commercial and financial networks. The efficient running of the global economy was based on the confidence in the effective working—and integrity—of the various institutions that comprised the system infrastructure (broadly defined as transportation, communications and financial services). No less importantly, it depended also upon the smooth movement around the world of cargo, information, and capital. The Great Recession of 2008 appears broadly to confirm that, when confidence in information and institutions slips, the world economy is thrown into turmoil and the welfare of billions of people around the globe is affected. For what is money but confidence in data and in the proper completion of the transactions built upon it? And what is a panic, or a recession (or depression), if not a collapse in business and consumer confidence?

The Idea of Economic Warfare

Toward the end of the nineteenth century, commentators with an interest in strategic affairs and political economy began to speculate that the ever-growing interdependencies and interconnections between the great industrial powers must reduce the likelihood of war between them. Such thoughts came less from idealism than from recognition of the implications that the newly created global economic system (and the infrastructure upon which it had been built) was remarkably fragile and sensitive to shock. This recognition produced widespread fear that a major war, even if confined to European powers, must dislocate the world economic system, destroying in its wake business confidence and thereby producing severe economic, political and social consequences around the world. "The future of war", Ivan Bloch wrote in his famous treatise on warfare, is "not fighting, but famine, not the slaying of men, but the bankruptcy of nations and the break-up of the whole social organization … In short, the economic results which must inevitably follow any great war in the present complex state of human civilization".

the global economy was based on confidence in transportation, communications and financial services

The financial panic of 1907 and attendant economic depression reinforced and heightened such ideas. The "Banker's Panic" in the United States precipitated an acute liquidity crisis, producing general economic disruption and a wave of business bankruptcies. In Britain, the noted economist Sir Robert Giffen was alarmed by the severity of the global financial storm, which he believed had exposed serious weaknesses in the foundations of modern finance. More than anything else, the panic demonstrated the danger of interdependencies within the new world economic system: a shock in one part of the world was certain to be transmitted instantaneously everywhere—the point being that economic crises would not be localized. What is more, commentators appreciated that on what might be termed an "explosively index", the 1907 crisis ranked low: a major European War would rank much higher. Yet even the relatively "non-explosive" 1907 panic caused widespread economic turmoil which brought in its train severe unemployment. If therefore followed that a larger explosion might lead to unemployment on a scale to cause social unrest and even political instability.

Recognition of the inherent fragility of the world financial system, combined with older ideas of dependency upon international trade, encouraged the belief that the "next war" must of economic and social necessity be short in duration. In the extreme, some argued that the dislocation to global trade, and the world economic system, would be so catastrophic as to raise the specter of social collapse—theoretically precluding the possibility at all of major conflict, an idea popularized by the Nobel Laureate Norman Angell. Although military planners rejected this extremist viewpoint, they nevertheless seem to have admitted the plausibility of the central argument, and many accepted that the primary issue in any future war would be not the military outcome but the implications for the globalized economy. Hence the remarkably widespread conviction in 1914 that the troops would be home before the leaves fell, or that the war would be over by Christmas—if victory could not quickly be achieved then a prompt negotiated peace would be necessary.

Standing at the epicenter of the global trading system, at the hub of the global communications network, Great Britain appeared to have more to lose than most in the event of such a financial and economic catastrophe. Yet within the Admiralty the prospect of a meltdown in the global economy appeared to offer Britain a unique strategic opportunity as well as a strategic danger. On the one hand, British naval planners, who naturally thought more about Britain's role in the global commons than did army planners, accepted that Britain certainly had a great deal to lose. On the other hand, they realized, so did others, especially Germany. From about 1901, Admiralty planners began toying with the strategic possibilities of harnessing their naval supremacy to Britain's effective monopoly control over the infrastructure of the international trading system in order to exacerbate economic derangement for the enemy while mitigating the impact upon their own economy. One might say, in effect, they contemplated "weaponizing" the global trading system.

Why did the Admiralty believe that Britain would suffer less from British efforts to derange the global economic system than would Germany? Basically because Britain's dependence on the smooth functioning of the global economic system was

matched by its control over the smooth functioning of the system, whereas Germany had no such control. In effect, Britain was in a position to deny Germany access to the system while retaining access for itself, for the following reasons:

- The Royal Navy was the most powerful navy in the world with an unrivalled capability to exert direct control over seaborne trade.
- The Admiralty possessed by far the most sophisticated information-intelligence gathering network in the world, and an understanding of how to lever this relative advantage into global situational awareness.
- Economically, Britain appeared better placed than any other nation to weather the financial and economic tsunami that was expected to strike at the outbreak of war—for instance bigger and better capitalized banks, a more flexible capital market, and above all creditworthiness. The state's ability in wartime to borrow and spend freely cannot be overstated. Problems could be solved and time literally 'bought' by the lavish spending of money.
- More importantly still, British companies dominated the physical and virtual infrastructure of the global trading system—the cable networks strung across the globe, the insurance and reinsurance industry centered in London, the banks financing international sales, the ships carrying commercial goods, etc. This conferred upon the British empire, or appeared to confer, the ability to wield a measure of control over the levers of international trade.
- Since the early 1900s, when strategic planners first began toying with the strategic possibilities that envisaged harnessing naval supremacy to Britain's effective monopoly control over the infrastructure of the international trading system, the Admiralty had forged very close links with a number of key companies, most notably Lloyds of London (insurance), Harris and Dixon (freight forwarding), and the Eastern Telegraph Company (cables). In modern parlance, steps were taken to develop "private-public partnerships" in the sharing of information.

Above all else, however, was a conceptual breakthrough—the realization that the strategic environment in which navies must plan and operate is substantially defined by the structure and character of the world economic system. The significance of this breakthrough should not be minimized: there is a difference between reality and the perception of reality, and there was no guarantee that the latter would catch up to the former.

The changing perceptions of naval planners unsurprisingly provoked strong resistance. The immediate response by the army's director of military operations, upon first learning of the navy's proposed new strategy, was to tell his opposite number in the Admiralty that the navy and the army apparently had "a very grave divergence of opinion … not so much on the general question of strategy as upon the whole question of war policy, if not indeed upon the question of what war means." This general did not exaggerate: the navy was indeed coming to a fundamentally different understanding of what war meant. He promptly illustrated the gap in understanding by labeling economic warfare "an invertebrate means of offence"—a wonderful turn of phrase! What he apparently meant was that war is not proper war, and armed forces are spineless ("inverterbrate"), unless they "break stuff". What the general missed was that the navy did, in fact, possess a spine—and, more to the point, possessed a much better sense of economic anatomy.

In particular, the navy recognized that the nervous and circulatory system of the global economy increasingly depended on the sea, though in ways, perhaps, that were not entirely obvious. Whereas sea "communications"—traditionally the target of naval pressure—had once been limited to ships carrying goods and letters, by the early twentieth century they encompassed:

- The communications grid—undersea cables and later wireless;
- The British steam merchant marine (as ever watched over by the Royal Navy);
- The networked international "financial services industry" centered in London—also the hub of the global communications grid—which allowed vendors to ship goods to purchasers on the guarantee that payment would be made.

It can be easy to miss the novelty and strategic significance of these developments. Put most simply, the flow of goods over the sea depended upon a parallel yet separate flow of information via cable. From the strategic perspective, furthermore, and of course employing an expanded definition of the term "communications", an array of new vulnerabilities—and opportunities—now existed. What is more, very little of this new expanded strategic environment was governed or regulated by internationally agreed rules and laws. Whereas there were plenty of precedents governing the interdiction of ships and goods in war-time, there were almost none governing the interdiction of electronic information. Yet seaborne trade could be interdicted just as well through non-naval as well as naval means.

The essence of the navy's proposed economic warfare strategy was for Britain to disrupt these "communications", through naval and non-naval means, so as to exploit the natural economic forces set in motion by the outbreak of war that were expected to cascade though the economies of all nations, leaving in their wake widespread financial and economic chaos. The idea was that Britain would take certain measures calculated to exacerbate and accelerate the derangement of the German economy with the aim of *quickly* collapsing Germany's underpinning systems and thereby provoking social upheaval. The aim of the strategy was not merely to interrupt enemy military operations but rather to dislocate civilian systems with the object of creating economic chaos and panic. The means were not to pressure choke points through simply restricting an enemy's maritime trade, nor precise attacks on specific individual industrial or military targets, but a wide range of actions designed to threaten confidence in the commercial access and financial systems underpinning Germany's economy. In other words, British actions were calculated to target both the physical and psychological. Weaponizing the infrastructure of global trade would translate into a shock—not attrition— attack upon an enemy society. The means and ends of this plan thus differed fundamentally from traditional blockade, which pre-dated the globalized economy, was predicated on outdated assumptions about the day-to-day conduct of international trade, and could work only slowly.

there were almost no precedents governing the interdiction of electronic information

Of course, however impressive these connections between "communications" and economic shock sounded in theory, theory was not reality. As they struggled to turn theory into a workable strategy, naval planners made and acted upon three key conceptual breakthroughs. First, they realized that they needed advice from the people who conducted and studied international trade if they wanted to understand the global economy in the necessary detail. As a result, they began speaking to leading economists, bankers, shippers, etc. Second, they realized that there would be very significant legal implications to British interference with global communications. As a result, Admiralty officials conducted what would now be called "lawfare," doing their best to ensure that British negotiators at international legal conferences such as the Hague Conference of 1907 favorably shaped the international maritime legal environment (though it must be said, with limited success, because the British plenipotentiaries would not cooperate). Third, they realized that interfering with communications was not an operational problem but a grand strategic problem, affecting the interests of multiple stakeholders, foreign and domestic, and requiring the highest political approval. As a result, they encouraged and participated energetically in inter-departmental discussions.

Victory in these inter-departmental discussions was far from assured. From the perspective of the political executive, the Admiralty's plan for economic warfare required revolutionary innovations in the strategic policy process and the assumption of enormous political risk. Both of these requirements derived from the extraordinary array of stakeholders whose interests would be affected by a campaign of economic warfare. They included British consumers, British businesses (especially in the shipping, communications, and financial services industries), and foreign neutrals. Within the British government, these stakeholders were represented chiefly by the Board of Trade, the Treasury, and the Foreign Office. Incidentally, the mere act of including the Board of Trade in strategic defense discussions was, by the standards of the day, revolutionary; traditionally, strategy had been a matter for the Admiralty, the War Office, and perhaps the Foreign Office. By the same token, the mere act of trying to enlist the support of British

business interests for the strategy—to say nothing of actually adopting or implementing the strategy—required substantial expenditures of political capital. The British government in the years before World War I was Liberal, which at that time meant an ideological commitment to free markets and free trade. Quite simply, seeking business support for wartime government control over the three pillars of global communications—telegraphs, merchant ships, and financial services—risked alienating the government's core constituency.

From February 1911 to May 1912, a group of senior government officials known as the Desart Committee met to assess the relative risks of economic warfare. Its establishment reflected the political executive's recognition that economic warfare was "too important a matter" to be left to the admirals: it was grand strategic rather than operational and that deciding upon it required input from multiple governmental and non-governmental stakeholders, not just from the Admiralty. The Desart Committee's investigation, which included testimony from leading bankers, shippers, and insurers, made clear that there would be powerful resistance from British business to economic warfare. The Board of Trade and the Foreign Office voiced their concerns about the domestic and foreign costs of the strategy very clearly. They argued, quite rightly, that economic warfare would entail large-scale state intervention in the workings of both the domestic and international economy, starkly challenging traditional ideas about the role of government. In so doing, moreover, it would far exceed established boundaries of what constituted grand strategy and indeed the very nature of war. The domestic and diplomatic fallout, they rightly predicted, was certain to be massive.

Though daunted, the political executive discounted these warnings. Impressed by the Desart Committee's assessments of the potentiality of economic warfare, the political executive (represented by the Prime Minister and eight other senior Cabinet ministers acting in conjunction with the Committee of Imperial Defence) gave the defense establishment permission to forge ahead with preparations for offensive warfare. The government resolved that in the event of war they would assert their right to intervene in the economy. In secret, the government drafted a set of regulations and penalties to govern the activities of British companies in wartime, so as to prevent their "trading with the enemy" or on the enemy's behalf. These were articulated in a series of Royal Proclamations, drafted pre-war, which forbade British merchants, financiers and shippers—indeed, any British subject throughout the empire—to trade or conduct business with the enemy. The naval and military authorities were further granted "pre-delegated authority" (another truly remarkable innovation in defense arrangements with huge constitutional implications) to implement immediately upon declaration of war stringent controls over a wide range of commercial enterprises connected with international trade.

By comparison with preparations to wage offensive economic warfare, defensive planning to survive wartime economic dislocation—some of which would be caused by Britain's own offensive measures—was minimal. Granted, as early as 1880 the British government began sponsoring the laying of "strategic" cables intended to build greater redundancy into the communications grid. In addition, the British government subsidized the shipping industry through awarding lucrative contracts to deliver the mail to distant parts of the world, and insisting that vessels used to carry the mail be new-built and exceptionally fast and thus be of use in time of war as auxiliary cruisers. During the 1900s, however, this approach was abandoned. Too much money had seemingly been wasted: further technological changes had rendered much of what had been built or done already obsolescent. Moreover, the British government had fallen upon fiscal hard times.

Efforts at defensive planning undertaken by the Desart Committee also faltered. A series of interviews with the leading bankers and insurers of the day left committee members in no doubt that the City of London regarded any government intervention (unless it was a backhanded subsidy) as an unwarranted and dangerous intrusion beyond the government's competence. The commercial interests refused to contemplate or discuss with the state any form of government economic intervention; they insisted that in time of war they would take any necessary defensive measures. And such alarms were amplified by the Treasury and Board of Trade. Consequently, intimidated by howls of protest at the high financial costs mooted, deterred by the intrinsic complexity of both the technology and dazzling interconnections within the system, and especially fearful of the high anticipated political costs entailed, the political executive declined to press the matter.

In retrospect, the actions—or inactions—of the political executive may seem foolish. If it dropped its defensive plans when confronted with resistance from the private sector, why did it adopt economic warfare, which required cooperation from the private sector? The likeliest answer is that they could see no better option. As the Desart Committee's investigation made clear, the Admiralty's plan for economic warfare was likely to cause huge collateral damage while it lasted—but it was not likely to last long. Many observers expected that any war would be short due to economic dislocation even without a deliberate campaign to worsen the dislocation. By contrast, the War Office's plan for the continental commitment offered far less credible hope for quick victory—and a prolonged war arguably carried even greater political risk than a short one, however brutal the short one might be. While it is reasonable to wonder why the political executive did not do more, it is more vital to understand the significance of what it did do. Even if the strategy of economic warfare was never implemented—after all, war in 1914 was not inevitable—actions taken in its formulation, adoption, and preparation would *and did* require the political executive to undertake revolutionary innovations in the policy process and to assume substantial political risk.

Implementation and Abandonment: August—October 1914

Britain declared war on Germany on 4th August 1914. Already the optimized global economic system had demonstrated itself to be highly sensitive to shock. The mere expectation of war during the last week of July had caused a virtual cessation of world trade, an impact even more dramatic than the most pessimistic commentator had imagined. Commodity exchanges around the world shut their doors. In the financial world the panic was even worse than expected—on a scale greater even than occurred in 1929/31. By 31st July, every stock exchange around the world (including Wall Street) had shut its doors. There was a global liquidity crisis. Banks recalled their loans. Foreign exchange was simply unavailable—though on the grey markets in New York, sterling was selling for $6 (up from $4.86). In London, meanwhile, the City was technically bankrupt; the accepting houses that funded international trade were unable to meet their obligations. The British government was compelled to step in, underwriting the entire stock of outstanding bills of exchange (in the world) and in so doing increasing the national debt obligation overnight by approximately 75%. The British then compounded this chaos by implementing their economic warfare measures—albeit only briefly. Although it is practically impossible to unravel the relative impact of the one or the other, it is clear that British implementation worsened the chaos—as had been intended.

Then, as also foreseen, a backlash arrived—but far more swiftly and intensely than expected. As the scale of the economic

devastation wrought became increasingly apparent, domestic interest groups became ever more vocal in clamoring for relief and lobbying for special exceptions, and neutrals howled in outrage at collateral damage to their interests. Within the government, their protests received a sympathetic hearing from officials at the Treasury, Board of Trade, and Foreign Office, who had never fully approved of economic warfare in the first place. Inadequate economic data clouded understanding and spawned uncertainty leading to hesitation; political commitment to the strategy began to crumble; and more and more exceptions to the published rules were granted, undermining the effectiveness of economic warfare. Implementation stalled.

In October 1914, aware of evasions and growing outright defiance by domestic interests, combined with mounting pressure from powerful neutrals (like the Woodrow Wilson administration), the economic warfare strategy was aborted. As a result, the British were compelled to wage war in ways they had previously agreed were undesirable, unthinkable, unworkable and even fatal. The reasons bear consideration by any nation contemplating similar warfare.

What Went Wrong?

The war exposed the limits of prewar planning—and of the political will to engage in prewar planning—in several ways. One showed the potential of relatively narrow technical details in economic warfare to have large political consequences. For example, in 1914, in order to prevent trading with the enemy, the British cable censors required that all messages be transmitted in plain English (i.e. no shorthand or abbreviations or code) and that each telegram include the recipient's full name and address of recipient (i.e. no internationally registered abbreviations such as LAMP32—akin to a dot com address like shoes.com. In so doing, however, they either forgot or did not appreciate the finite limitation in cable capacity—or "bandwidth", to use the modern term. The effect of the new regulations on telegram content doubled or even tripled the length of each message. The result was a communications log-jam and commercial paralysis. Overruling objections from the military censor, the government quickly relented, dialed down the regulations, and agreed to share communications resources with corporations (both nominally British and foreign).

Another way in which the war exposed the limits of prewar planning concerned the behavior of British businesses. Before the war, faced with abundant evidence that they would resist government regulation but seeking to avoid a politically damaging confrontation, the government defaulted to blithe hopes about private-sector conduct. These included the expectation that moral suasion would translate into effective control, that businesses would cooperate with regulations, and that capitalists would forgo enormous opportunities to make profit on the black market out of patriotism. Such an assumption ignores the reality that capitalistic economies are built upon a 'reward system' that encourages firms (and individual businessmen) to deviate from the conventional and pioneer new methods: those who succeed earn disproportionate rewards; those who fail risk bankruptcy. Put crudely, the instinctive and essentially rational behavior of businessmen is to make money through innovative means. It might be said that conforming to government expectations is antithetical to the business mentality.

Aside from the political costs of confrontation, the structure of British business made measuring its compliance with regulations extremely difficult. Tracking large corporations was one thing; tracking small businesses, through whom an enormous amount of economic activity flowed, was another. Generally speaking, there exists an inherent conceptual bias when talking about the problem of envisioning the economy in terms of large corporations, big systems and big data. In reality a vast (un-quantified) amount of economic activity flows through the enormous base of small business. In any case, the point here is that the pre-war British government never set up sufficient detection and enforcement mechanisms to ensure compliance with the announced prohibitions on trade.

As a result, certainly within six months, perhaps within three, British banks were financing most of the contraband trade from Americans to Germany via neutrals, deals with Germany were being transacted over British cables, and the goods carried to the enemy in British ships. Although these violations were apparent to some degree, the military authorities responsible for waging economic warfare found themselves powerless to prevent these absurdities. Early attempts to improvise a better organization were resisted by other government departments

(whose assistance was needed), while political leaders turned a blind eye. In the meantime British trade with previously unknown corporate entities located in countries contiguous to Germany grew exponentially. For more than a year the British government remained unaware of the scale of the problem, lacking the means to gauge it, and not wanting to believe the worst.

The government's ability to impose effective control over the economy developed only gradually, and not because British businessmen suddenly discovered a hidden reservoir of patriotism. By 1916, many in the private sector were sufficiently worried by the prospect of a social revolution that they were willing, it would seem, to tolerate relatively moderate state interference as a preferable alternative to arbitrary confiscation of private property by a radicalized "Socialistic" society. The government and businesses had different understandings of what constituted a security emergency: for the government, the national security emergency was the prospect of military defeat; while for businesses, the corporate security emergency was the prospect of social revolution. In other words, when businesses finally began to cooperate with the government, they did so not because the government's prewar expectations about corporate patriotism were correct, but because they came to fear something more than government regulation.

The government's failure to anticipate the behavior of British businesses reflected an even more fundamental failure to reach consensus with key stakeholders about the proper relationship between state and society in wartime. While the authority of the state to conscript its citizens was well-enough established during the nineteenth century, cemented by Prussia's victory over France in 1870, the state's right to conscript never extended to private property. Social cooperation with a strategy that affected property interests had to be voluntary; it could not be legally compelled (and still cannot?). For the government, voluntarism was necessary not only to avoid legal challenge but to acquire the information—in effect the "targeting data"— needed to prosecute the strategy of economic warfare. National-security imperatives required society to reconceptualize its relationship to the state, but neither party realized the degree to which reconceptualization was necessary.

The Commons Strike Back

In seeking to disrupt global communications, broadly defined, via economic warfare, Britain enjoyed a number of advantages:

1. A near monopoly over the "communications" infrastructure of international trade, defined here as telegraphs, merchant shipping, and financial.
2. Naval officials with the imagination to understand that the character of the global economic system defined the navy's operating environment and to spot new strategic opportunities caused by changes in the global economic system.
3. Naval officials who acknowledged that they lacked expertise on the day-to-day conduct of international trade and were willing to seek assistance from economists, bankers, shippers, etc.
4. Naval and other officials who realized that any attempt to interfere with maritime communications posed serious legal problems and attempted to shape the legal terrain accordingly.
5. The broad recognition that any attempt to interfere with maritime communications was a grand strategic rather than an operational problem, requiring input from multiple stakeholders, inside and outside government.
6. A political executive willing to conduct strategic discussions with multiple stakeholders, even at the risk of alienating its core political constituency.
7. Strong prewar political commitment to the strategy of economic warfare, manifested concretely in the predelegation of authority.

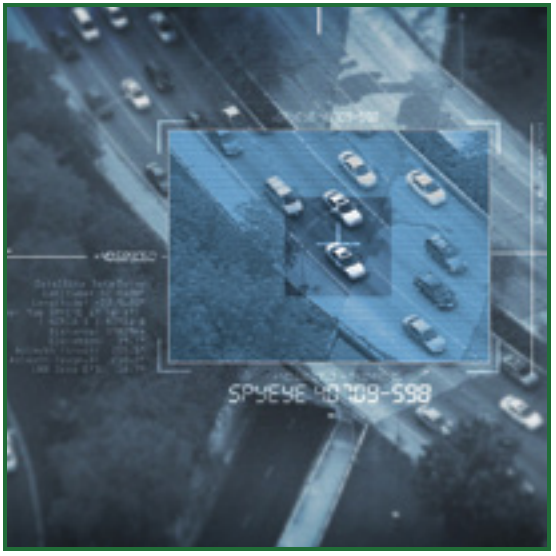
Even with all these advantages, Britain's strategy of economic warfare still failed—indeed it was barely tried.

The planners of cyber warfare could use this story to assure themselves that they would not make the same mistakes today. Or they could use it as an opportunity to ask whether the United States is in the same situation as Britain, and to think through some of the obstacles they might face today.

One obstacle is simply to define cyberspace. Just as definitions of maritime "communications" were not self-evident before World War I, so definitions of cyberspace are not self-evident.

How does one distinguish private cyberspace from public cyberspace, and American cyberspace from foreign cyberspace? Will it suffice to defend just U.S. military cyber systems and U.S. critical infrastructure? Surely the U.S. economy, which depends on access to cyberspace, must be regarded as a critical interest? Further, given that the health of the national economy depends so very greatly upon a healthy world economy, should not national security measures encompass this too? Where does one draw the line?

Alternatively, it may be helpful to think through potential parallels between maritime space and cyberspace. The maritime space most readily identified as the global "commons" are the "high seas," or oceans. But they are contiguous with progressively smaller and more sovereign (i.e., non-common) waters—gulfs, bays, deltas, ports and harbors, rivers, inland seas, lakes, etc.—some of which may be reachable by continuous voyage, some of which may be more isolated. Determining exactly where a commons turns into a sovereign area is not easy. By analogy, the private/sovereign areas of cyberspace are the cyber equivalents of inland seas, ports and harbors, great lakes, etc., public/foreign access to which may require permission from, or the use of force against, a host. The idea of cyberspace as a commons co-exists uneasily with various private/sovereign claims.



As if the challenge of defining interests in cyberspace is not enough, it also involves very difficult legal questions—just as the challenge of defining maritime "communications" posed very difficult legal questions. The fact is that U.S. firms dominate cyberspace, and very large portions of global Internet traffic pass through the United States. Does it follow that the U.S. government can legitimately claim the right to commandeer or withhold bandwidth by which one gains access to cyberspace—in other words, the radio frequencies used by the mobile communications systems through which users will increasingly interact with cyberspace, as mobile platforms and related "apps" proliferate? Similarly, the U.S. government, without necessarily "owning" cyberspace, presumes to control access to portions known as .gov and .mil—but that is because they have control over the servers and the content. Do they own the interaction space, the protocols and gateways by which people gain access to those servers and data? Does exercising sovereignty amount to legal ownership? Does "ownership", per se, confer the right to control access? Does the U.S. government have a legal or moral right to defend, regulate or control access to cyberspace in ways that will very likely impinge on others' interests? In time of war these questions will push their way to the front of political awareness. The exact answers are less important here than the recognition that these questions must be asked. The British experience with economic warfare suggests that it would be very dangerous for the U.S. government to *assume* that it could readily translate national dominance of cyberspace—however defined—into legal or effective state control.

If it does decide to wage cyber warfare, how will the United States insulate itself from the collateral damage caused by deranging the global commons? How will the U.S. government gain the cooperation and monitor the compliance of American companies? How will it respond when American consumers complain about rising prices, and when American businesses protest heavy state regulation, unfair foreign competition and falling profits? What will it do when allies and neutrals complain that actions targeted at belligerents might as well be targeted at them? The issue is not necessarily that the United States will be unable or unwilling to act unilaterally within cyberspace. Rather, it is that IF the United States acts unilaterally, THEN, for a variety of reasons—primarily economic but also political and diplomatic, not to mention legal—it will impinge on the critical interests of others and face a severe backlash. Effective measures that the United States

might take in cyberspace could hurt American and foreign interests so much that the U.S. might be compelled to call off its attack, just as the UK had to do in October 1914.

If the British experience with economic warfare has a single lesson, it is that the infrastructure of a globalized economic system makes for a weapon of mass disruption rather than a precision weapon. Accordingly, weaponizing it is a grand strategic, pervasively political problem. It is not a problem for computer experts or the Pentagon alone. To have any hope of success, a strategy to weaponize critical economic infrastructure requires the recognition that multiple stakeholders—foreign and domestic, inside and outside the government—exist, and it requires efforts to gain their cooperation. Its formulation demands direction from the highest political authority and the assumption of substantial political risk by elected officials even to seek cooperation from powerful constituencies, let alone to alienate them by actually implementing the strategy. The more aggressive the weaponization of the global economic infrastructure, the more severe the damage it will cause, not only to its intended target but to "collateral" stakeholders, including neutral nations, domestic business interests, and domestic consumers who vote. For the strategy to survive the likely backlash, or for the intensity of the backlash to be reduced, a case must be made to stakeholders *before* the strategy is implemented that the costs of an alternative strategy—or no strategy—would be even worse—say, a war that drags on for four years, costs millions of lives, and raises the specter of revolution at home. It may be impossible to secure the cooperation of all interested parties, but it is certainly impossible to do so without realizing that their cooperation is necessary.

In the event of a future major conflict, waging economic warfare within the context of a very different global economic structure would, as it did a century ago, be quite different in its character than anything experienced before. It thus behooves us, now, to devote serious and persistent thinking to the subject. ❇

NOTES

1 This paper began life as a report prepared for a U.S. CyberCom Project on Cyber Analogies. I have removed footnotes in the belief that doing so was appropriate for this forum, but interested readers can find my evidence and further historical explication in *Planning Armageddon: British Economic Warfare and the First World War* (Cambridge: Harvard University Press, 2012).

Silicon Valley: Metaphor for Cybersecurity, Key to Understanding Innovation War

John Kao[1]
Chairman, Institute for Large-Scale Innovation

Recent attention has focused on the value of metaphors for deepening our understanding of cybersecurity and defining useful pathways for action. Metaphors carry two types of value for such work. First, the utility of a metaphor lies in its ability to help us understand a complex phenomenon, to create a gestalt, or a narrative, that can expand the available space for creative thinking. That in turn shapes the second form of utility, namely enhancing our ability to frame useful responses, whether in the form of prototypes, experiments, initiatives or policies. Therefore my guiding questions in this paper are three: How does Silicon Valley as a metaphor increase our knowledge of cybersecurity?; What are the implications for action?; and How do these in turn contribute to our larger understanding of the phenomenon of innovation war?

Silicon Valley has received widespread attention, and if imitation is the sincerest form of flattery, then one need look no further than the website http://tbtf.com/siliconia.html to see the proliferation of other geographies that emulate Silicon Valley's example, at least in name. This website currently lists close to a hundred geographies with the word "silicon" in the title, hence Silicon Glen (Scotland), Silicon Wadi (Israel), Silicon Alley (New York), Silicon Plateau (Bangalore) and even Siliwood (Hollywood).

This level of homage is not accidental or arbitrary. Silicon Valley can rightfully claim today to be the world's leading innovation epicenter, or "habitat", as some of its denizens prefer to call it. Consider just one tentpole of the Silicon Valley phenomenon—Stanford University—which has been the birthplace for more than 6,000 new ventures. They read like a who's who of the innovation economy. Companies said to have been developed during a student or researcher's time at Stanford: Atheros Communications, Charles Schwab & Company, Cisco Systems, Cypress Semiconductor, Dolby Laboratories, eBay, e*Trade, Electronic Arts, Gap, Google, Hewlett-Packard, IDEO, Instagram, Intuit, Intuitive Surgical, Kiva, LinkedIn, Logitech, MathWorks, MIPS Technologies, Nanosolar Inc., Netflix, Nike, NVIDIA, Odwalla, Orbitz, Rambus, Silicon Graphics, Sun Microsystems, SunPower, Taiwan Semiconductor, Tensilica, Tesla Motors, Varian, VMware, Yahoo!, and Zillow.[2]

Or consider venture capital, another essential ingredient of the Silicon Valley model. Fully 40% of U.S. venture capital—funds with current assets of some $11 billion—are located in Silicon Valley, many on a single street, the fabled Sand Hill Road.[3]

Other elements of Silicon Valley's storied innovation ecosystem include a past history of significant U.S. government funding, additional academic institutions such as UC Berkeley, a host of entrepreneurship savvy service providers—law firms, banks, accountants, human resources consultants etc.—as well as high quality lifestyle and weather.

The result, in terms of wealth generation, is evident; as one marker, Bay Area median household income, including data from four counties not directly related to Silicon Valley, was 41% higher than the U.S. national average and 37% higher than the rest of California.[4]

THE SILICON VALLEY SUCCESS MODEL—RISK-TAKING AND OUT-OF-THE-BOX IDEAS

What are some key aspects of the Silicon Valley business model?

First, Silicon Valley has a culture of risk-taking that is unique worldwide. Investments can be made on the strength of a conversation or a sketch on the back of a napkin. Venture

capitalists enter into deals knowing that the majority of them will fail. Angel investors deploy their capital based on "well-honed intuition."

Such an attitude toward risk mirrors American culture writ large. The American comfort with risk-taking and wild ideas remains our singular, absolute advantage in a world that is increasingly filled with the raw materials for innovation—talent, good ideas and the infrastructure to support them. It is still the case that countries that talk the talk of innovation, for the most part, fail to extend full support to the lonely risk-taker who is light on success guarantees. Resources may be forthcoming, provided that failure does not rear its ugly head. Even in countries that are seemingly as innovation-friendly as Singapore and Finland, business failure is viewed with a dim eye. Likely as not, you will be ostracized from polite business society, refused business loans and other niceties.

In contrast, the Silicon Valley saying is "if you haven't gone bankrupt at least once or twice, you're not trying hard enough." A favorite investor question to an entrepreneur in Silicon Valley is to "describe your biggest failure." The correct answer is not to minimize your failures, but rather to come up with a major "crash and burn" story. A seasoned Silicon Valley venture capitalist will understand that failure is inextricably woven into the entrepreneurial process and is much more interested in how an entrepreneur dealt with failure than whether it occurred.

Venture capitalists enter into deals knowing that the majority of them will fail.

This accords with psychological explorations of risk-taking that involve the well known ring-toss exercise. Psychologists have studied entrepreneurs based on where they place themselves in relation to the ring-toss post. If the entrepreneur stands too close to maximize the likelihood of success, that is not an encouraging sign of a willingness to try new unproven things. Conversely if they stand too far away and lob the rings at the post, they are assuming an excessive amount of risk. So finding the "sweet spot" for risk becomes the order of the day.

Related to risk-taking is tolerance of off-the-wall ideas, what Howard Rheingold, noted social media expert, calls "wild ass ideas." The evocative film, *Something Ventured*,[5] portrays titans of the Silicon Valley venture capital community talking about Apple Computer. One after another, they admit that they missed the "next big thing," perhaps put off by the eccentricities of Steve Jobs and Steve Wozniak, thereby losing a claim to turn an initial investment of $350,000 into a current market capitalization of some $200 billion. This highlights the inherent unpredictability of the innovation process. While incumbents may be allergic to this level of uncertainty, the ability to see things that others do not, to follow hunches and to sense useful innovation from the edges is key to innovation success.

Genuine risk capital and the attitudes that come with it are fundamental to the Silicon Valley phenomenon. The classic model of Silicon Valley venture capital marries a visionary investor with a technology visionary to create a new industry, as has been the case over and over again with companies like Tandem Computers, Amgen, Apple, Intel, Google, and many others. Nowhere else in the world is it so possible to get started with a simple conversation and handshake.

Venture investors, to be clear, are not simply shoveling money randomly out the back of a truck, but rather are using disciplined intuition coupled with a sophisticated internal assessment of investment-grade entrepreneurs to guide their decisions. While capital is abundant in many other places in the world, it resembles banking money rather than genuine early-stage risk capital. There is an old adage: "You can't get money from a bank until you don't need it." This certainly is the case with what passes for venture capital in many other parts of the world.

Silicon Valley is distinct in the extent to which it offers early-stage money for ideas, and its tolerance for the iteration and experimentation that an unproven idea must go through to have proof of concept validated. This dynamic—now captured in the notion of the lean startup[6]—is intrinsic to the innovation process. New ideas have to be tried out and field-tested. They cannot simply be asserted as revealed truth. Business approaches that emphasize planning and analytics for an early-stage idea often produce a mismatch of expectations. There

is really no way that a business plan can be valid in an early-stage environment. Like the old adage of "no plan survives first contact with the enemy," no business plan survives first contact with the market. The point here is that sophisticated strategic conversations are key to a startup, but five-year pro formas and business plans may be less relevant. It is therefore easy to understand why new ideas can become implemented in the Silicon Valley environment with tremendous speed and flexibility.

THE SILICON VALLEY SUCCESS MODEL—COMMUNITY

Less obvious but no less important to the Silicon Valley phenomenon is the importance of community. I have compared Silicon Valley and its startups to the time-honored tradition of a barn-raising. In the pioneer days, if you were a newcomer to town, the community would gather and give you their collective effort to build a barn and at the end of the day you'd have one. Your obligation then reciprocally would be to offer a day's worth of labor to the next new person who required it and so on. The rough and tumble of Silicon Valley competition, which perhaps had its origins in the California Gold Rush of the 19th century, is more than made up for by the extraordinary efforts of those who have made their fortunes in Silicon Valley to give back to their community with their time and money to endow professorships, support community development programs, and even fund a home-town museum, The Tech Museum of Innovation in San Jose.

This notion of barn-raising is very much part of the hidden DNA of Silicon Valley. There are few other communities lubricated with as much reciprocity and smooth networking. There is a kind of specialized emotional intelligence at work. People know what's going on, who to make the proper introduction to, who's in the club, not in the club, who would make a good partner, ally, friend. As a result, the Silicon Valley dynamic is not typically a linear progression from idea to funding to execution, but rather the expression of a web of relationships, an ecosystem in which there are many pathways to yes, and in which the scarce knowledge of how to practice the innovation process is sprinkled across many different kinds of professionals, whether they are lawyers, venture capitalists, or even the waiter at Buck's Restaurant, the famous gathering place in Silicon Valley, who may in fact have a good idea of his or her own.

This sense of community certainly lies at the heart of the university, business and entrepreneurial community nexus in Silicon Valley and the often cozy business-academic ties that result. It is well-known, for instance, that John Hennessy, president of Stanford, also gave Sergey Brin and Larry Page resources that helped with the launch of Google and subsequently received equity and a board seat in that company. This behavior might be viewed with skepticism by American colleagues and with suspicion by peers in other countries, despite the fact that communities as diverse as Helsinki's Aalto University and Moscow's Skolkovo Foundation are trying to emulate these kinds of town-gown relationships in growing their own versions of Silicon Valley.

THE NEW GEOGRAPHY OF INNOVATION AND THE RISE OF COMPETITION

As Silicon Valley has matured over the past half-century, much has been made of its growing pains. Concerns about the future of Silicon Valley have been expressed in such terms as the high cost of residential real estate that freezes out young creators, the fevered bidding wars for talent that increase the cost of starting up, the arrogance and complacency that comes from mold-breaking levels of success.

But the real issue for Silicon Valley is globalization and the proliferation of innovation capability worldwide. How can Silicon Valley maintain its preeminence in an "innovation world?"

It is worth reviewing some economic history. At the conclusion of World War II, the United States was really the only game in town when it came to innovation. And Silicon Valley in its early days was a singular phenomenon. America had economic might, an intact homeland, universities that the world's talent clamored to come to, an R&D base second to none, and so on. It is striking to recall that, post-World War II, the United States at its high water mark of economic and social influence

was responsible for no less than 50 percent of global economic output. Clearly this was also a time in which America's innovation prowess was at an all-time high and driven by a host of absolute advantages.

But now, innovation capabilities have gone global and all of America's advantages, save its culturally defined comfort with risk, mentioned above, have become relative. The world is striving to become competent at the kind of societal innovation that is increasingly seen as driving economic growth and social development. I have elsewhere coined the term large-scale innovation[7] to refer to innovation at a societal as opposed to enterprise scale. Most of our knowledge about innovation comes from studying companies, while the study of innovation in societies remains a great work to be done.

If one adopts a large-scale innovation perspective, it is clear that there are at least 50 countries around the world with sophisticated national innovation agendas, strategies, and investment programs. China alone is putting $500 billion into its national innovation agenda, while the government of Sweden for example has Vinnova, an agency with 300 professionals dedicated to pursuing a national innovation strategy. Helsinki has created an Innovation University by merging the faculties of engineering, business, and design at their leading universities. And governments ranging from Russia to Singapore are investing in large-scale innovation platforms. Russia has the Skolkovo Foundation that explicitly intends to emulate the success of Silicon Valley with the support of American multinationals such as Intel and Cisco. Singapore, a country with no natural resources save smart people, targeted life sciences as an area in which to achieve eminence and now boasts the Singapore Biopolis, a life sciences city within a city that is on track to having 6,000 PhD scientists working on its premises.

In this "innovation world," it is clear that new ideas can come from anywhere. And multinational corporations are playing their roles in enabling that. For example, Cisco has a Network Academy education effort that counts a million students across the globe. The efforts of China to link access to domestic markets with the establishment of multinational corporate laboratories in China is well known. In a speech at the Office of the Director of National Intelligence a few years ago, I commented that the labs of today are the foxholes of tomorrow, and that this dissemination of innovation capability can be likened to a swarm of innovations and innovators that will inevitably lead to unexpected breakthroughs from unexpected quarters. Welcome to innovation world.

In military doctrine, the swarm references the power of the many and the small to fight the large and centralized. So, for example, ultra-fast gunboats loaded with explosives may go up against aircraft carriers. There is a similar phenomenon in innovation. The centralized corporate labs of today and yesterday, the Bell Labs for example, are increasingly confronted by innovation coming from the margins, from a kind of "innovation swarm" made of up of venture-backed companies and "super-empowered" innovators for whom the cost structure of doing their innovation work may be a fraction of what it used to be.

Disruptive innovation, the kind of innovation that changes the nature of the game, typically does not come from incumbents but rather from insurgents who are crowding in from the margins and may lack traditional credentials but whose ideas—perhaps outlandish at first—may wind up setting the new standards. It is often the odd idea that can create a new order of things. Who can forget how Xerox with its attachment to the mainframe model of copying (proprietarily called Xerography) fell victim to Canon with its expertise in miniaturization and optics that capitalized on a market for personal copying made up of entrepreneurs, the

self-employed, and community organizations? Xerox was very happy serving its current market while Canon perceived a new one, and the rest is history, as Xerox suffered a near-death experience. Consider, therefore, box cutters and passenger airplanes or from a cybersecurity perspective the many thousands of military- or kleptocracy-grade hackers working to overturn the existing order of things in bunkers, basements and faceless office buildings around the world.

SO WHAT FOR CYBERSECURITY?

Innovation capability is now becoming abundant and global. The old adage of brainstorming is that quantity leads to quality. Inevitably, amassing innovation capability over time around the globe will lead to breakthroughs that may come from any point of the compass. In a way, the world itself is becoming more like Silicon Valley, with disruptive innovation able to originate from virtually everywhere. The cybersecurity community would therefore be well advised to understand the Silicon Valley model with an eye to enhancing its own innovation capability and also perceiving clearly the new competitive landscape within which Silicon Valley is no longer a singular phenomenon.

Bringing Silicon Valley inside will be a challenge for military institutions that are used to hierarchy, elimination of uncertainty, authority, centralization of resources and top-down mechanisms of control. Recall that early-stage initiatives typically cannot be validated by traditional methods of proof and documentation—business plans and financial models may not capture the essence of a new idea or validate it. Cutting off outlier ideas is not, however, a smart strategy, if there is any reasonable chance that they might show promise in the uncertain landscape that is innovation war. The insurgents meanwhile—our cyber warfare adversaries—are all about outlier ideas and asymmetric responses.

So the cybersecurity community would do well to ponder the ingredients of the Silicon Valley model if it wishes to go up against the "swarm" that sits on the other side of the battlelines in the kind of amorphous innovation war that we are now engaged in. This swarm is leaderless, crowd-sourced, crowd-funded, emergent, experimental and, like Silicon Valley, is able to reap the benefits of adopting a different kind of business culture from that of the mainstream.

So the cybersecurity community might usefully address some core questions inspired by the Silicon Valley example:

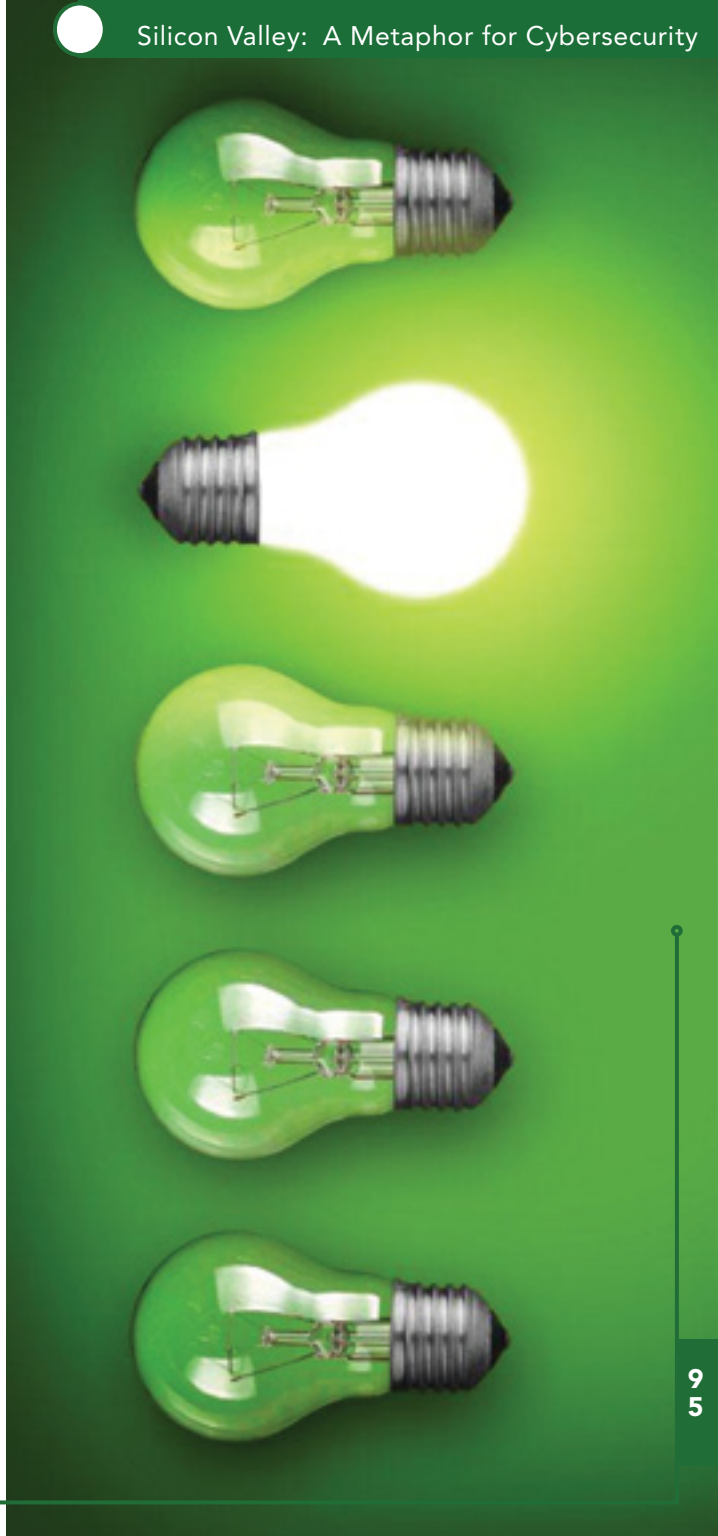
1. How can it build enduring capabilities analogous to leading universities to ensure the continuing development of deep knowledge about salient agendas?
2. How can it practice a model of risk capital and investment that does not require the usual validation, but instead supports emergent possibilities in a manner similar to the early-stage venture capital community of Silicon Valley?
3. How can it create an attractive environment for talent to work in a fluid and collaborative way that is free from inhibiting influences of hierarchy and the day-to-day, and that instead supports the kind of "white space" most conducive to the creative thought that leads to disruptive innovation?
4. How can it win the war for talent by offering attractive opportunities to do one's best work? How will it create the kind of "strange attractors" to draw talent in by focusing on cool ideas, horizon opportunities, and the overall mission?
5. How will talent attraction in turn lead to talent multiplication in terms of achieving critical mass or "share of talent?"
6. How can it create the kind of trust, collaboration and community internally that creates salience, speed and support? Will the good guys be able to make an innovation model out of an ability to validate emergent ideas on the back of a napkin through a network of trust?

Overall, the cybersecurity community will have to learn how to build relationships across the hard skin of organizational boundaries, to be able to establish friendships, alliances, partnerships and collaborations that don't fit the traditional model of defense contracting. In addition, security classification becomes an issue, because many of the kinds of people that one might most wish to learn from either are not eligible for clearance, do not want it, or would regard the notion of classification as too onerous.

In conclusion, it is important to understand that to fight the swarm, we have to be the swarm; to fight the innovation swarm, we have to be the innovation swarm. The intelligence community needs not only to study and interact with Silicon Valley, but to figure out what it wants to apply from the Silicon Valley playbook. Keeping pace with the "long war" is actually the struggle to fight the innovation war, within which the building of innovation capabilities globally has created new instability as well as new collaborative opportunities. ❇

NOTES

1 john@johnkao.com
2 Piscione, Deborah, Secrets of Silicon Valley, Palgrave, 2013
3 Florida, Richard, The Geography of Venture Capital, Atlantic Online, January 25, 2012
4 Bay Area, A Regional Economic Assessment, Bay Area Council Economic Institute report, October 2012
5 Miralan Productions, Something Ventured; risk, rewards and the original venture capitalists, distributed by Zeitgeist Video
6 Blank, Steve, *The Startup Owners Manual,* self-published, Amazon
7 largescaleinnovation.org



The Offense-Defense Balance and Cyber Warfare

Keir Lieber

Offense appears to trump defense in cyberspace. Cyber attacks are cheap, whereas cyber defense is expensive. Defenders must defend across an entire network, whereas attackers merely need to find individual points of vulnerability. Cyber attack is considered technically easier and can be carried out quickly, whereas cyber defense is complicated and time-consuming. Because of the difficulties of attributing the source of a cyber attack, attackers can expect to hit and live to fight again another day. And because modern military command and control systems are so dependent on information technology, cyber attacks offer the tantalizing prospect of strategic decapitation. In contrast, strategic victory through pure cyber defense seems remote.[1]

This observation has led some analysts to predict a new era of international cyber instability and conflict. Specifically, scholars and policymakers worry that cyber offensive advantage will provoke spirals of hostility, arms racing, and the outbreak of full-fledged cyberwars.

This paper argues that such concerns are overblown. The conventional wisdom is wrong for two reasons: First, the offense-defense balance analogy itself is conceptually flawed and empirically unsupported. Second, regardless of any merits, the offense-defense balance concept travels poorly to the realm of cyberspace. In particular, the inherent complexity and ambiguity surrounding cyberwar preparations and operations precludes generating the kind of state behavior predicted by offense-defense theory. In short, potential U.S. adversaries are unlikely to see greater incentives to launch cyber attacks to degrade U.S. military capabilities than they otherwise would in the absence of perceptions of a cyber offensive advantage.

> many of the steps pursued by states to bolster their own security make other states less secure

The paper is organized as follows. The first section presents the basic tenets and predictions of offense-defense theory, the body of academic work that underpins modern concerns about offense-dominance in cyberspace. The second section discusses why offense is thought to trump defense in cyberspace, and why such a finding points to potentially troubling consequences for U.S. national security. Third, the paper analyzes how these problems are exacerbated when the offense-defense balance analogy is applied to cyberspace. Finally, the paper concludes with policy recommendations.

OFFENSE-DEFENSE THEORY AND THE PROSPECTS FOR CONFLICT

Contemporary concerns about the consequences of cyber offense dominance rest on a body of academic work known as "offense-defense theory" (ODT). The core claim of the theory is that the nature of technology at any given time is an important cause of international conflict: leaders will be more tempted to launch wars when they believe new innovations favor attackers over defenders.

ODT is perhaps best known from the passionate and intricate debates about nuclear first-strike capability and deterrence stability during the Cold War, as well as from the popular interpretation—made famous by historians Barbara Tuchman *(The Guns of August)* and A. J. P. Taylor *(War by Time-Table* and other works)—that World War I was triggered by erroneous perceptions of a new era of swift and decisive warfare.[2] The theory remains a staple of international relations theorizing, and it shapes modern policy debates on arms control, national security, and defense force structure and posture. Below I discuss the theory's background, the core concept of the offense-defense balance, and how the offense-defence balance (ODB) analogy has been used to explain several key historical cases.

Background

Allusions to the relative strength of attack and defense, and the idea that offensive advantages foster conflict, whereas defensive advantages promote peace, can be traced back to

the writings of Sun Tzu, Carl von Clausewitz, and Antoine-Henri Jomini. For example, Clausewitz wrote that "the greater strength of the defensive [might] tame the elemental fury of war"; after all, he continued, "if the attack were the stronger form … no one would want to do anything but attack."[3] In the last century, basic offense-defense insights are found in the analytical work of B. H. Liddell Hart and J. F. C. Fuller in the 1930s; Marion William Boggs and Quincy Wright in the 1940s; and Malcolm Hoag and Thomas Schelling in the 1960s.[4]

Specialized works focused on ODT, however, were produced in the 1970s and 1980s at a time when Cold War strategic nuclear issues loomed large. International relations scholars such as George Quester, Robert Jervis, Stephen Van Evera, and Charles Glaser believed that nuclear weapons gave defenders a large military advantage, ensured strategic stability in the form of mutual deterrence, and should have allowed U.S. and Soviet leaders to feel tremendously secure and act more peacefully. The superpower nuclear arms race confounded these views, leading offense-defense proponents to conclude that U.S. and Soviet leaders misunderstood the nature of the nuclear revolution. And proponents of ODT feared that the unnecessary arms competition could spiral into an otherwise avoidable nuclear war.

In this context, the originators of ODT sought to identify how military technology in general might cause war or peace.[5] The seminal work in this vein is Jervis' 1978 article, "Cooperation under the Security Dilemma." The security dilemma holds that many of the steps pursued by states to bolster their own security simultaneously make other states less secure.[6] Even when such preparations are meant solely for self-defense, other states are compelled to interpret these military preparations as hostile and prepare accordingly, because international anarchy (the absence of a central authority to govern or protect states), pervasive uncertainty about present or future intentions, and the shifting balance of power make inaction or the attribution of benign intent inordinately risky. This triggers an action-reaction spiral of reciprocal arms buildups, diplomatic tension, and hostility that can lead to war.

The severity of the security dilemma depends on whether offense or defense has the advantage on the battlefield (i.e., ODB) and whether offensive and defensive capabilities can be distinguished.[7] When the balance favors offense, the probability of competition and war is greater for several reasons. Striking first seems an attractive option for all states. If technology is such that states believe an initial attack will lead to a quick and decisive victory, then striking first will be tempting—both because of fear of attack under such conditions and the desire to make gains through conquest. Second, political crises are more likely to escalate into full-fledged military conflicts because states will be quicker to conclude they are threatened, more inclined to attack preemptively, and more prone to inadvertent wars.[8] Third, arms races become more intense because even small advantages in armament levels can have decisive consequences in a short war. In sum, all types of war—e.g., expansionist, preemptive, preventive, accidental—are more likely when the balance favors offense. The opposite holds when defense has the advantage.

The Offense-Defense Balance

The ODB turns out to be a surprisingly complex concept to operationalize in practice. In order to apply the ODB to cyberspace, it is necessary to first clarify some basic definitions and assumptions:

The ODB is typically defined as the relative ease of attack and defense, with "relative ease" denoting some measure of the relative costs and benefits of attack and defense. "Offense" and "defense" refer to the use of military force, not the political motivations, intentions, or goals that may motivate such military action. Specifically, offense is identified as the use of force against another state in order to seize territory or destroy assets. Defense entails the use of force to block those attacking forces. The relative cost of attack and defense is understood as the amount of resources (money) an attacker must invest in offensive capability to offset the amount of resources a defender invests in defensive forces.[9]

Offense and defense can be used to describe military action at three levels of warfare: strategy (pertaining to overall war plans and ultimate outcomes), operations (the conduct of specific campaigns in a war), and tactics (concerning actions and engagements taken within a campaign or battle). In order to measure the ODB and its effects on military and political

outcomes, the strategic level would appear to be the most suitable frame for analysis. That is, because the theory ultimately aims to explain the politics of war and peace (i.e., decisions to initiate war), what matters most are leaders' expectations of whether attackers can win wars quickly and decisively. After all, launching a war would not be appealing to leaders if the nature of the balance promised victory in the opening campaigns but defeat in the war. In practice, however, the operational level of warfare offers a suitable unit of analysis, especially for measuring the objective ODB. For one thing, the success or failure of specific campaigns will very often determine whether the attacker or defender prevails in a given war. And, given the great uncertainty military and political leaders face in predicting the course of any war very far into the future, the prospect that initial campaigns can be won rapidly serves as a good functional equivalent for the prospect of strategic success. In sum, the logic of offense-defense theory implies that the strategic ODB is the key unit of analysis, but in practice the ODB at the level of operations is what should drive military and political outcomes.

The ODB should be distinguished independently of great disparities in power and skill between adversaries. Battlefield outcomes clearly depend on a host of factors other than the ODB, including the overall balance of military forces, resources, and skill. If a given state wins a quick and decisive victory, it could stem from the nature of the ODB, but it could also simply result from a gross disparity in military power (understood in quantitative or qualitative terms). To classify an ODB requires the assumption that states make reasonably intelligent decisions about how to employ existing technologies and forces given prevailing knowledge at the time.

Finally, one needs to stipulate the criteria used to judge how technology favors offense or defense at any given time. Scholars have struggled to identify objective and consistent criteria, but at least two hypotheses emerge: First, mobility-improving innovations are presumed to favor attackers because they contribute to quicker and more decisive warfare. Second, firepower-improving innovations seem to favor defenders because they result in longer, more indecisive warfare.

Historical Cases

Scholars have used offense-defense logic to explain the history of major wars, ethnic civil wars, arms races, alliance behavior, crisis behavior, military doctrines, and much more. Consider, for example, how the ODB analogy has been used to explain the origins of the Wars of German Unification, World War I, World War II, and the Cold War arms race:

Wars of German Unification

The spread of railroads in the mid-1800s—the quintessential technological innovation of the industrial age—dramatically increased the strategic mobility of armies. The mobilization, deployment, and concentration of ever larger forces could now be achieved across vast distances at up to ten times the speed of marching troops. Offense-defense theory predicts that the greater mobility conferred by railroads made quick and decisive victories for the attacker more likely. Moreover, the prospect of quick victory offered by railroads (and the fear that an adversary would seize the same opportunity) should have made leaders more inclined to launch war. Prussia's quick and decisive victories in the Wars of German Unification—against Denmark (1864), Austria (1866), and France (1870–71)—appear to validate the theory's predictions of railroad-based offensive advantage. Although the outcome of the Danish War was a foregone conclusion given across-the-board power disparities, Prussia's victories over Austria and, especially, France were surprising, given that both states appeared to have greater overall military strength and fighting experience as well as extensive rail networks. What else could explain how upstart Prussia could invade and defeat in short order the two recognized great powers of the continent besides the role that railroads played in benefiting the attacker? Moreover, according to the story, Prussian leaders recognized and opportunistically embraced the offensive advantage of railroads in order to pursue an expansionist foreign policy.

World War I

The Great War is the paradigmatic case for the ODB analogy. Scholars contend that the dramatic increase in firepower created by the revolution in small arms and artillery in the late 19th and early 20th centuries gave an enormous advantage to defenders and resulted in longer and more indecisive warfare.

Tragically, however, European statesmen and military leaders erroneously perceived great offensive-dominance in the succession of new firepower technologies. In turn, these massive misperceptions of the ODB—the "cult of the offensive" and "short-war illusion," as they have been labeled—served as a master cause of World War I. As A. J. P. Taylor wrote, "When cut down to essentials, the sole cause for the outbreak of war in 1914 was the… belief in speed and the offensive."[10] As Robert Jervis describes it, "A version of spiral dynamics was… an immediate cause of the outbreak of war in 1914. Each of the continental powers believed that the side that struck first would gain a major military advantage. Since to wait for the other side to clarify its intentions could mean defeat, even a country that preferred the status quo to a war would feel great pressures to attack…[11] Had the participants recognized the objective defensive advantages of the new weapons of war, Jervis writes elsewhere, "they would have rushed for their own trenches rather than for the enemy's territory."[12] The "cult of the offensive" dominated German planning in particular, because military and civilian leaders were mesmerized by "a highly exaggerated faith in the efficacy of offensive military strategies and tactics."[13] Stephen Van Evera writes, " between 1890 and 1914 Europeans increasingly believed that attackers would hold the advantage on the battlefield and that wars would be short and decisive. They largely overlooked the lessons of the American Civil War, the Russo-Turkish War of 1877–78, the Boer War, and the Russo-Japanese War, which had revealed the power of the new defensive technologies."[14] Van Evera concludes, "Europeans embraced political and military myths that obscured the defender's advantages… and primed Europeans to expect a quick, total victory for the stronger side in the next war."[15]

World War II

The mechanization and motorization of armies—especially in the form of the tank—transformed land warfare between the world wars, greatly increasing operational mobility on the battlefield. Proponents of offense-defense theory believe that the incorporation of tanks into the European armed forces in the interwar period resulted in greater offense-dominance and provoked Adolf Hitler's decision to attack his enemies. According to Stephen Van Evera, "During 1919–45 the power of the offense was restored by motorized armor and an offensive doctrine—blitzkrieg—for its employment."[16] Robert Jervis notes, "the German invasion in World War II… indicated the offensive superiority of highly mechanized armies in the field."[17] Sean Lynn-Jones writes, "The tank, for example, is useful for offensive and defensive purposes, but without tanks, blitzkrieg offensives would be virtually impossible. Tanks make it possible for states to launch offensives using large armored formations. In other words, they make offensive strategies far less costly than they would have been without tanks."[18] And Adolf Hitler, because of his recognition that armored warfare gave offense a great advantage over defense, felt free to pursue his expansionist aims. To be sure, Hitler had very aggressive aims in any case; but the ODB analogy suggests that he was truly emboldened only when he grasped the offensive promise of armored blitzkrieg warfare.

Cold War Arms Race

The invention of nuclear weapons produced an unprecedented, exponential increase in firepower—and the equivalent of a revolution in defensive advantage. "After 1945 thermonuclear weapons restored the power of the defense, this time giving it an overwhelming advantage," Van Evera writes. According to Glaser, "Nuclear weapons created a revolution for defense advantage."[19] Thus, according to scholars, the nuclear revolution rendered strategic nuclear competition between the United States and the Soviet Union irrelevant, irrational, and dangerous.[20] Nuclear weapons should have essentially eliminated the security dilemma and much of the grounds for security competition. Yet, as with the case of World War I, leaders failed

to appreciate the true nature of the ODB, and undertook an intense and costly nuclear arms race—involving highly offensive strategic forces—that threatened to precipitate World War III. In the end, and despite rampant misperception, the fact that the Cold War never turned hot is seen as a testament to the ultimate power of defense embodied in nuclear weapons technology.

In sum, the analytical power of the ODB analogy is that it conceivably explains why all kinds of states—aggressive or defensive, revisionist or status quo, greedy or insecure—might face incentives to attack based simply on the nature of the technology available at the time. World War I was a tragic accident—a war that nobody really intended, but one that was unleashed by rampant misperceptions of the new industrial tools of warfare; World War II was driven in large part by Hitler's discovery of blitzkrieg warfare, which gave him the offensive firepower to pursue his boundless objectives; and the Cold War nuclear arms race was an illogical product of leaders who failed to appreciate the revolutionary defensive (or at least deterrent) character of the absolute weapon. Based on the ODT literature, it seems reasonable to worry that the current era of cyber offensive advantage could spell trouble.

A world replete with cyberwars appears to loom ahead.

But is offense-defense theory useful for understanding cyber threats to U.S. national security? Is the ODB analogy—the idea that conflict is more likely when offense dominates—applicable to cyber warfare? In short, does the cyber ODB foretell a new era of instability and conflict?

THE OFFENSE-DEFENSE BALANCE IN CYBERSPACE

John Arquilla speculates that cyberwar could be "showing the world that outbreaks of conflict may be primarily driven by the state of play in technology." "Today," he continues, "this state of play is one that makes attacking seem easy and defending oneself hard. A world replete with cyberwars appears to loom ahead."[21] To evaluate this and related claims, one must first operationalize the ODB in the context of cyberspace. How can we understand the relationship between offense and defense in cyberspace? Why is offense thought to be dominant in cyberspace? What predictions flow from such a finding?

War is a political act—its purpose is to achieve political objectives. In an anarchical world fraught with uncertainty, states are compelled to care about one objective above all others: national security. Offense-defense theory purports to explain when states are more or less likely to resort to war to achieve national security goals. According to the theory, offensive advantage in military operations makes all kinds of wars more likely: expansionist wars driven by states seeking to expand their power over other states; preventive wars driven by states seeking to stop the rise of a potentially threatening rivals; preemptive wars driven by first-strike incentives and fears; and unintended wars sparked by spirals of arms racing and hostility.

If the ODB analogy is to shed any light on the dynamics of cyberspace, the analysis should be restricted to the realm of cyberwar. That is, the realm consisting of military operations conducted in cyberspace to destroy, damage, or degrade an enemy's military capability in order to achieve political ends.

To be sure, analysts and policymakers—including those at U.S. Cyber Command—are interested in a larger set of cyber threats. Broadly defined, cyber attacks also encompass non-strategic attacks aimed at manipulating, stealing, disrupting, denying, degrading, or destroying critical systems, assets, information or functions. But if the ODB analogy is to have any analytical purchase in cyberspace, it cannot be conceptually stretched to account for all of these cases. Put simply, offense-defense theory is all about explaining why and when technology encourages states to attack each other. Even if the ODB analogy were not applicable to crime, terrorism, espionage, or any number of other cyber operations, its application to understanding strategic cyberwar among states would be an important and useful accomplishment.

Cyber Offense Dominance

Although it is widely agreed that offense has the advantage over defense in cyberspace, few studies have explicitly adopted the

offense-defense framework to rigorously examine the claim. Nevertheless, both existing and new criteria point toward a clear offensive advantage in cyberspace.

Mobility

According to the traditional criteria discussed above, mobility favors offense because it multiplies an attacker's advantage in surprise and initiative. That is, greater mobility increases an attacker's ability to quickly outflank or overwhelm a surprised defender and reduces the time an attacker must take to assault defensive positions. Cyber offense is highly "mobile" in the sense that attacks can be carried out almost instantaneously. Cyber defenders may have little (if any) warning of a cyber attack, may not even detect an attack as it is underway and inflicting damage, and may be quickly outflanked or overwhelmed. In short, for cyber attackers, achieving surprise is easy, targets can be assaulted immediately, and operations can be completed very quickly. Mobility in cyberspace appears to favor offense in spades.

Cost

Measures of the relative ease of attack and defense typically rest on relative costs—the amount of resources an attacker must invest in offensive capability to offset the amount of resources a defender invests in defensive forces. Determining such an actual cost ratio in a real-world case is enormously difficult, if not impossible, but it seems safe to say that such a ratio in cyberspace favors the attacker. One analyst—using a heuristic model combining empirical data and case studies on the offensive and defensive costs of hardware, software, and personnel—estimates that the offense-defense ratio in cyberspace favors attackers over defenders by 132:1.[22] Another analyst calculated that a high-end "cyber army" capable of defeating U.S. cyber defenses could be developed for roughly $100 million dollars, which pales in comparison to the annual U.S. cyber defense budget.[23] In short, the cost of defending computer networks appears to be far more expensive than penetrating those networks.

> the offense-defense ratio in cyberspace favors attackers over defenders by 132:1

Complexity

Much of the cost discrepancy between cyber offense and defense stems from the relative technical difficulty of these operations. In large part because defenders must defend across an entire network, whereas attackers merely need to find individual points of vulnerability, cyber attack is considered technically less demanding.[24] Cyber attackers also appear to require fewer personnel and less computer science knowledge to prepare and launch attacks than defenders require to prevent such attacks. According to the 2010 U.S. Quadrennial Defense Review, an offensive advantage "is growing as hacker tools become cheaper and easier to employ by adversaries whose skills are growing in sophistication."[25] As Deputy Defense Secretary William Lynn put it, "A couple dozen talented programmers wearing flip-flops and drinking Red Bull can do a lot of damage."[26]

Other Attributes

Other attributes of cyberspace appear to favor the attacker even if there does not appear to be a clear corollary in traditional offense-defense theory. First, because of the difficulties of attributing the source of a cyber attack, attackers can expect to strike and live to fight again another day. Second, improved knowledge of defensive cyber operations, tools, and strategies can simultaneously be a huge boon to attackers. Finally, because modern military command and control systems are so dependent on information technology, cyber attacks offer the tantalizing prospect of strategic decapitation. In contrast, strategic victory through pure cyber defense seems remote.

Predictions Based on Cyber Offensive Advantage

Offense dominance in cyber space has led many to identify an emerging "cybersecurity dilemma" that will generate a world replete with cyber arms racing, spirals of hostility, and the outbreak of full-fledged cyberwars.[27] The United States, in particular, might be dragged into a range of cyber conflicts. For example, committed adversaries—those that want to harm the United States and damage U.S. capabilities—could

seize the opportunity created by offensive advantage to launch significant cyber attacks against U.S. military assets. Potential adversaries—those that have conflicts of interest with the United States but are otherwise deterred from taking hostile actions—might be emboldened by offense dominance to pursue revisionist goals through cyberspace, or launch attacks for fear of the U.S. striking first. Even the United States, as it continues to build up its offensive cyber capabilities, might be tempted to launch more preventive or preemptive cyber attacks in the name of national security.[28]

The ODB analogy strongly suggests that cyber offense dominance would exacerbate the security dilemma and lead to more arms racing, and conflicts driven by preemptive, preventive, and aggressive incentives:

Arms Racing

As Lord and Sharp write, "Offensive dominance creates a great risk of cyber arms races. States… are likely to view the prevalence of offensive cyber threats as a legitimate rationale for bolstering their own capabilities, both defensive and offensive, thus fueling an action-reaction dynamic of iterative arming. Experts believe that at least 20 nations are engaged in a cyber arms competition and possess the type of advanced capabilities needed to wage cyberwar against the United States. As Michael Nacht, Former Assistant Secretary of Defense for Global Strategic Affairs, told us, 'An arms race is already going on in cyberspace and it is very intense.'[29] Such arms races are predicted to result in the outbreak of war because they trigger an escalating spiral of hostility and misunderstanding. As Lord and Sharp continue, "Conflict in cyberspace is uniquely predisposed to escalation given uncertainties about what constitutes an act of war and the growing number of state… actors seeking offensive capabilities. Actors are more likely to misperceive or miscalculate actions in cyberspace, where there is no widely understood strategic language for signaling intent, capability and resolve. Uncertainty will encourage states to prepare for worst-case contingencies, a condition that could fuel escalation."

Preemptive War

The ODB analogy also suggests that countries will face greater incentives for taking preemptive military action. If attack is relatively easy, the incentive to strike first is great, because a successful surprise attack provides larger rewards and averts looming dangers. If aggressors and potential adversaries are persistently vulnerable to cyber attack, there is little hope for the emergence of stability, predictability, and trust.[30] Instead, states—facing few if any observable signs of imminent attack, the prospect of an attack being carried out instantaneously, and the danger of surprise strategic decapitation—will be sorely tempted to land their own decisive blow first. In short, as states build more and more offensive cyber capabilities, they will grow more trigger-happy (mouse-happy?)—launching preemptive cyber attacks to exploit the advantage of the initiative and deny the same opportunity to the adversary.

Preventive War

Cyber offense might also encourage states to launch more preventive cyber attacks. From an ODT perspective, the development and deployment of the Stuxnet worm to attack Iran's nuclear enrichment facilities was a harbinger of preventive cyberwar—marking a dangerous turning point where states will be tempted to reach for new and increasingly disruptive offensive cyber weaponry to address or eliminate military threats before they fully materialize. The fact that diplomatic or international legal restrictions have not advanced in lockstep with these capabilities means that there is little to check the strategic temptation to mount preventive attacks.

Wars of Aggression

Finally, according to the ODB analogy, cyber offense-dominance creates a greater likelihood of direct attack by a belligerent adversary. In particular, analysts worry about the possibility of a surprise attack designed to strike at the heart of country's military power. Just as the right combination of offensive technologies and tactics emboldened Adolf Hitler to launch quick and decisive blitzkrieg operations against France in 1940, adversaries might seize upon offensive cyber weapons to attack the United States without warning and with devastating strategic consequences. Former U.S. Secretary of Defense Leon Panetta labeled this danger a "cyber-Pearl Harbor," with cyber attacks being used in conjunction with physical military attacks to cripple infrastructure, and kill and traumatize Americans.[31] Offensive cyber attacks could be so decisive

that, according to Andrew Krepinevich, "much of the military capability of the United States could prove to be the modern equivalent of the Maginot Line." Other analysts echo Krepinevich's fear of a terrible surprise attack that would neutralize U.S. military advantages.[32]

PROBLEMS WITH APPLYING OFFENSE-DEFENSE THEORY TO CYBERSPACE

Despite its prominence in international relations scholarship, potential policy relevance, and intuitive appeal, offense-defense theory is deeply flawed. The core ODB concept is almost impossible to define, operationalize, and measure in any meaningful way. And the purported behavioral consequences stemming from state perceptions of the ODB are empirically unsupported. These problems have been discussed at length elsewhere, and need not be repeated here.[33]

Applying these concepts in the cyber realm is equally problematic. The most general reason to doubt the utility of the ODB analogy in cyberspace stems from enormous uncertainty about whether cyber attacks can inflict significant military damage on a victim. Such uncertainty undermines the applicability of ODT predictions—about arms spirals, opportunistic wars of aggression, preventive wars, and preemptive wars—to cyberspace.

Offense-defense theory is based on the idea that offense-dominance makes quick and decisive victory possible, which in turn makes initiating war more attractive under a range of circumstances. Just as German leaders allegedly based their strategies for quick and decisive victory in Europe on perceived offense-dominance before World War I and II (with the Schlieffen Plan and Blitzkrieg, respectively), so too might countries opt for cyber attacks to deliver quick victory in the future. But, as Thomas Mahnken notes, "Cyber advocates have failed to offer a theory of victory for cyberwar, a chain of causal logic linking the use of cyber means to the achievement of political ends."[34] It is reasonable to question whether cyber attacks can deliver victory in future conflicts, even in a supporting role with cyber attacks being used to bolster traditional military operations. It is even harder to see states (especially U.S. adversaries) being tempted enough by the offensive possibilities of cyberwar to initiate conflict.

We have little evidence so far of cyber weapons being used to cause physical damage, and almost none to military hardware. Clearly, the most likely kinds of cyber attack in the military realm would target U.S. C3I and logistical networks. But these sorts of attacks do not inflict any direct lethal damage, and their ability to either significantly undermine kinetic military operations or do so in a way that would contribute to successfully defeating or coercing an adversary is questionable. Russia deployed cyber attacks against Georgia in support of military operations in 2008, but it is hard to extrapolate from this example to future attacks by weaker powers against the United States.

Again, the power of the ODB analogy is that it helps explain why and when states that would otherwise be content or forced to live with the status quo would decide to launch military attacks. The key logic is the appeal of winning quick and decisive victory. Cyber attacks, at present, appear to lack the prospect of having such decisive military and political effects. Cyber attacks have a limited shelf-life and require follow-up kinetic attacks to have any hope of military impact. And, in any event, such attacks are unlikely to dramatically shift the overall conventional military balance between states.

Perhaps most importantly, attackers would face a great deal of uncertainty in predicting the strategic effectiveness of cyber attacks. Such uncertainty is inherent to the enterprise—not only because of the difficulty of predicting how one's own cyber attacks will develop and spread, but also because of great uncertainty about a target's defensive capabilities, ability to identify the attack quickly, and the ability to respond to such attacks. Preparatory probes and preliminary attacks are more likely simply to reveal vulnerabilities to the defender, which can then address them, than help the attacker discover weak points for subsequent exploitation. In short, it is hard to see how cyber attacks could be launched with the kind of confidence about effectiveness that underpins offense-defense logic. Consider each of the main ODT predictions:

Arms Racing/Spirals of Hostility

The ODB analogy about the increased danger of arms racing under offense-dominance seems ill suited to understanding the dynamics in cyberspace. To be sure, the United States is ramping up its cybersecurity forces in response to evolving threats. But most of our understanding of evolving cyber threats comes from evidence of actual attacks; i.e., from ongoing offensive operations. The logic of the security dilemma, by contrast, is that states living in a world of offense-dominance are forced to respond to military preparations by another actor, even if those preparations are for defensive purposes only: State A builds arms. State B sees those steps and is compelled to build more of its own arms to restore the military balance. State A sees State B's response, fears B's intentions, and so builds yet more arms. This spiral of reciprocal arms buildups generates such hostility and suspicion that both States are tempted to launch an attack during a crisis.

In cyberspace, however, one country would have difficulty detecting that another country is ramping up its cyber capabilities in the absence of actual ongoing cyber attacks. If State A is not launching cyber attacks against State B, then State B is unlikely to perceive that State A is undertaking preparations for cyber attack—and therefore unlikely to respond nervously with preparations of its own, which then contribute to the outbreak of conflict. In other words, the kind of preparations and counter-preparations that drive otherwise security-seeking states into an unintended and unstable arms race seem absent in cyberspace. If one state is already engaged in cyber attacks against another, there is no security dilemma dynamic caused by offense dominance. There is simply one state pursuing aggressive aims against another.

Preemptive War

Preemptive strike incentives are vastly reduced when actors are uncertain of the effects of military attack. According to the logic of ODT, it is the *clarity* of effects—the decisive advantage of the first mover—that can lead states to attack regardless of whether they have aggressive or defensive aims. But that kind of clarity about effects is absent in cyberspace. Would a future adversary have any high degree of confidence that its preemptive cyber attacks would succeed as planned? Would the adversary (especially a relatively weaker one) bank its entire strategy of follow-up military operations on the ability to achieve decisive advantages through a cyber first-strike?

Note that once a state is engaged in conventional combat, uncertainty about military effects is much less of an obstacle to action. States simply try various methods and learn from trial and error. But even in the conventional world, history does not provide many examples of true preemptive strikes.[35] And the main reason for this is a lack of confidence that a surprise first strike would give the attacker a decisive advantage—especially when weighed against the cost of starting a war. The crisis of confidence seems even worse in cyberspace.

Preventive War

Preventive attacks are launched for the purpose of stalling the growth of an adversary. The typical scenario is of a state trying to stem its decline in relative power by using force against a rising state while that state is still vulnerable—that is, strangling the baby in the crib before it grows up to harm you. But preventive attack can also be more appealing for a rising state if offense is dominant and the rising state fears being strangled in the crib.[36]

How does this logic play out in cyberspace? If a rising China, for example, sought to undermine U.S. power before that power could be used to stifle China's growth, it would be far more likely to employ other tools of statecraft. For example, China might resort to financial actions aimed at wrecking the U.S. economy. Moreover, China has already built—and continues to build—powerful asymmetric conventional military options that it could use to undermine U.S. power if it felt threatened by U.S. behavior. China has not resorted to either of these courses of action because it is deterred. Or it has determined that such steps would be counterproductive or simply ineffective. Cyber capabilities seem unlikely to change this basic calculus because China would face great uncertainty about the impact of cyber attacks, especially about medium- and long-term effects.

Wars of Aggression

The most promising case for applying the ODB analogy to cyberspace might lay in the argument that cyber offense-dominance will embolden a belligerent adversary to launch a direct attack. An aggressive adversary might be tempted to launch a surprise attack if it thought it could strike decisively at the heart of its enemy's military power. But the "cyber Pearl Harbor" scenario is far-fetched. Objective analysts have raised much skepticism about the ability to conduct truly decisive offensive cyber warfare, and states are unlikely to be more optimistic than armchair analysts. The uncertainty behind Japan's gamble in 1941 was less about its ability to inflict damage on the U.S. fleet at Pearl Harbor—and thus on the U.S. ability to project power in the Pacific—than about the will of U.S. leaders and the American people to sustain support for a long and bloody conflict. (It is also worth noting that few states in the future are likely to face such desperate circumstances as those faced by Japan in 1941.) By comparison, uncertainty about both the military effectiveness of a cyber bolt-from-the-blue and the political impact of such an attack would seem to overwhelm any theory of decisive strategic victory.

In short, the inherent uncertainty about the effects of offensive operations in cyberspace undermines predictions derived from the ODB analogy. States that would otherwise accept the status quo are unlikely to be led by cyber offense down a spiral of hostility into arms racing and preemptive or preventive attacks. And there is no compelling reason to believe that hostile U.S. adversaries will face greater incentives to launch cyber attacks than they otherwise might in the absence of any comprehension of a cyber offense advantage.

CONCLUSION

This paper examines what offense-dominance in cyberspace implies for U.S. Cyber Command's understanding and articulation of its role and mission in U.S. national security policy. On the surface, there seem good reasons for concern. Cyber attacks have clear advantages over cyber defense. According to offense-defense theory, offensive advantage in military operations exacerbates the security dilemma among states, leading to a greater likelihood of dangerous arms races and attacks driven by preemptive, preventive, and aggressive motives. To make matters worse, and in comparison with the conventional realm of operations, cyberspace entails rapid technological innovation and much uncertainty about the nature and source of attacks. All of this suggests great cause for worry.

However, the answer offered here is that although U.S. commanders should be concerned about many cyber threats, the prospect of greater conflict stemming from offense dominance is not one of them. In particular, instability and competition among states that would otherwise be willing to accept the status quo seems highly unlikely. To be sure, offense is relatively easier than defense in cyberspace. But the inability to sufficiently predict and control the consequences of cyber attacks serves as a major check on the incentive to resort to such attacks. Neither fearful nor aggressive states will find an answer to their national security problems in the cyber realm.

The equivalent of a "null" research finding might initially seem unsatisfying for policymakers and military planners seeking to make sense of the dynamics of cyberspace. But identifying exaggerated security fears is a crucial step in formulating wise strategy. U.S. military planners need to devote cyber resources to deal with the threats that matter, not those that don't. Moreover, U.S. policymakers need to marshal persuasive counterarguments to any claims that the growth of U.S. cyberpower—especially in the realm of active defenses and offensive capabilities—is somehow dangerous and destabilizing.

The United States should maintain and bolster its impressive cyber capabilities because doing so can only strengthen deterrence. Relatively weaker states are unlikely to be emboldened by the offense-dominant nature of cyber weapons. The possibility of reaping important military advantages, much less winning quick and decisive victories, by resorting to cyber attack is at present remote. U.S. adversaries that are sophisticated enough to develop and deploy effective cyber capabilities are unlikely to be deceived by such a possibility, particularly if they face a real and credible threat of retaliation. ❇

NOTES

1 The idea that cyberspace greatly favors the offense is widely accepted, including among U.S. government officials and agencies. For example, see Department of Defense, *Quadrennial Defense Review Report* (Washington D.C.: U.S. Department of Defense, 2010), p. 37; and statements by General Keith Alexander, Commander, United States Cyber Command, including testimony before the Senate Armed Services Committee, April 15, 2010. For more in-depth assessments of the relative strength of cyber offense and defense, see Adam P. Liff, "Cyberwar: A New 'Absolute Weapon'? The Proliferation of Cyberwarfare Capabilities and Interstate War," *Journal of Strategic Studies*, Vol. 35, No. 3 (June 2012), pp. 401–428, at pp. 414–417; Andrew Krepinevich, *Cyber Warfare: A "Nuclear Option"?* (Washington, D.C.: Center for Strategic and Budgetary Assessments, 2012); Nicholas C. Rueter, "The Cybersecurity Dilemma," Thesis, Department of Political Science, Duke University, 2011; Ned Moran, "A Historical Perspective on the Cybersecurity Dilemma," *Insecure Magazine*, Issue 21 (June 2009), pp. 112–116; and Kristin M. Lord and Travis Sharp, eds., *America's Cyber Future: Security and Prosperity in the Information Age* (Washington, D.C.: Center for a New American Security, June 2011), pp. 20–31; and Matthew D. Crosston, "World Gone Cyber MAD: How 'Mutually Assured Debilitation' is the Best Hope for Cyber Deterrence," *Strategic Studies Quarterly*, Vol. 5, No. 1 (Spring 2011).

2 Barbara W. Tuchman, *The Guns of August* (New York: Macmillan, 1962); A. J. P. Taylor, *War by Time-Table: How the First World War Began* (London: Macdonald, 1969); and A. J. P. Taylor, *The First World War: An Illustrated History* (1963; repr., New York: Perigree Books, 1980). Also see B. H. Liddell Hart, *History of the First World War* (London: Faber, 1972); and L. C. F. Turner, *Origins of the First World War* (New York: W. W. Norton, 1970).

3 Carl von Clausewitz, *On War*, ed. and trans. Michael Howard and Peter Paret (Princeton: Princeton University Press, 1976), pp. 217–218, 359. On offense-defense ideas in Sun Tzu and Jomini, see Stephen Duane Biddle, "The Determinants of Offensiveness and Defensiveness in Conventional Land Warfare," Ph.D. diss., Harvard University, 1992, pp. 22–30.

4 J. F. C. Fuller, "What Is an Aggressive Weapon?" *English Review* 54 (June 1932): pp. 601–605; B. H. Liddell Hart, "Aggression and the Problem of Weapons," *English Review* 55 (July 1932): pp. 71-78, especially pp. 72-73; B. H. Liddell Hart, *The Memoirs of Captain Liddell Hart* (London: Cassell & Company, 1965), p. 186; Marion William Boggs, *Attempts To Define and Limit "Aggressive" Armament in Diplomacy and Strategy*, University of Missouri Studies, Vol. 16, No. 1 (Columbia: University of Missouri, 1941); Quincy Wright, *A Study of War* (Chicago: University of Chicago Press, 1965), pp. 807–810, 1518–1521 (the first edition of this book was published in 1942); Hoag, "On Stability in Deterrent Races"; and Schelling, *Arms and Influence*, pp. 224–225, 234–235.

5 See Robert Jervis, "Cooperation under the Security Dilemma," *World Politics* 30 (January 1978), pp. 167–214; Quester, *Offense and Defense in the International System* (New York: John Wiley & Sons, 1977), pp. 7, 208; Stephen Van Evera, *Causes of War: Power and the Roots of Conflict* (Ithaca: Cornell University Press, 1999); Jack Snyder, "Civil-Military Relations and the Cult of the Offensive"; Shai Feldman, *Israeli Nuclear Deterrence: A Strategy for the 1980s* (New York: Columbia University Press, 1982), pp. 45–49; Stephen Van Evera, "The Cult of the Offensive and the Origins of the First World War"; and Charles L. Glaser, *Analyzing Strategic Nuclear Policy* (Princeton: Princeton University Press, 1990).

6 Jervis, "Cooperation under the Security Dilemma," p. 169. The concept of the security dilemma is present, but less developed, in John H. Herz, "Idealist Internationalism and the Security Dilemma," *World Politics*, Vol. 2 (January 1950): pp. 157–180; and Herbert Butterfield, *History and Human Relations* (London: Collins, 1951). Charles Glaser reviews, clarifies, and extends the logic of Jervis' article in Charles L. Glaser, "The Security Dilemma Revisited," *World Politics*, Vol. 50 (October 1997): pp. 171–201.

7 The following discussion draws primarily on Jervis, "Cooperation under the Security Dilemma," pp. 187–199; and Glaser, "Security Dilemma Revisited," pp. 185–187.

8 This echoes George Quester, who writes: "If both sides are primed to reap advantages by pushing into each other's territory, war may be extremely likely whenever political crisis erupts. If the defense holds the advantage, by contrast, each side in a crisis will probably wait a little longer, in hopes that the others will foolishly take the offensive." Quester, *Offense and Defense in the International System*, p. 7. On the dangers of first-strike advantages, see Thomas C. Schelling and Morton H. Halperin, *Strategy and Arms Control* (New York: Twentieth Century Fund, 1961), pp. 14–16; and Thomas C. Schelling, *The Strategy of Conflict* (New York: Oxford University Press, 1963), chap. 9.

9 An offense-defense cost ratio may be unobservable in practice and in principle. This issue is addressed below.

10 Taylor, *War by Timetable,* p. 121.

11 Jervis, *Perception and Misperception in International Politics,* p. 94.

12 Jervis, "Cooperation under the Security Dilemma," p. 191.

13 Stephen Van Evera, "Why Cooperation Failed in 1914," *World Politics,* Vol. 38, No. 1 (October 1985), pp. 80–117, at p. 81.

14 Van Evera, *Causes of War,* p. 194. Van Evera notes, "Most German officers and civilians thought they could win a spectacular, decisive victory if they struck at the right moment." Ibid, p. 204.

15 Van Evera, *Causes of War,* p. 194.

16 Van Evera, *Causes of War,* p. 162.

17 Jervis, "Cooperation under the Security Dilemma," p. 197.

18 Sean M. Lynn-Jones, "Offense-Defense Theory and Its Critics," *Security Studies*, Vol. 4 (Summer 1995): pp. 660-691 at p. 676.

19 Van Evera, *Causes of War*, p. 162; Charles L. Glaser, "When Are Arms Races Dangerous? Rational versus Suboptimal Arming," *International Security*, Vol. 28 (Spring 2004), pp. 44–84, at p. 75.

20 Quester, *Offense and Defense*; chap. 13; Jervis, "Cooperation under the Security Dilemma"; Robert Jervis, *The Illogic of American Nuclear Strategy* (Ithaca: Cornell University Press, 1984); Jack Snyder, *The Ideology of the Offensive: Military Decision Making and the Disasters of 1914* (Ithaca: Cornell University Press, 1984); Robert Jervis, *The Meaning of the Nuclear Revolution: Statecraft and the Prospect of Nuclear Armageddon* (Ithaca: Cornell University Press, 1989); Charles L. Glaser, *Analyzing Strategic Nuclear Policy*; Van Evera, *Causes of War*, chap. 8.

21 John Arquilla, "The Computer Mouse that Roared: Cyberwar in the Twenty-First Century," *Brown Journal of World Affairs*, Vol. 18, No. 1 (Fall/Winter 2011), pp. 39–48, at p. 39.

22 Patrick J. Malone, "Offense-Defense Balance in Cyberspace: A Proposed Model," Thesis, Naval Postgraduate School, 2012.

23 Glenn Chapman, "Two Years and 100 M Dollars Buys Winning Cyber Army," Agence France-Presse (1 August 2010), as cited in Lord and Sharp, *America's Cyber Future*, p. 28.

24 As cited in Lord and Sharp, America's Cyber Future, p. 28: "According to the Defense Advanced Research Project Agency (DARPA), the number of lines of code included in security software increased from several thousand 20 years ago to nearly 10 million today. Over the same period, the number of lines of code included in malware remained constant at approximately 125.120 In other words, cyber defenses have grown exponentially in effort and complexity, but they continue to be defeated by offenses that require far less investment by the attacker."

25 Department of Defense, *Quadrennial Defense Review Report* (Washington D.C.: U.S. Department of Defense, 2010), p. 37.

26 U.S. Department of Defense, Office of the Assistant Secretary of Defense (Public Affairs), "Remarks on Cyber at the RSA Conference As Delivered by William J. Lynn, III, San Francisco, California, Tuesday, February 15, 2011."

27 The dark specter of cyberwars generated by technologies that make attacking seem easy and defending hard is raised, for example, in John Arquilla, "The Computer Mouse that Roared: Cyberwar in the Twenty-First Century"; Lord and Sharp, *America's Cyber Future*, pp. 29–30; Moran, "A Historical Perspective on the Cybersecurity Dilemma"; and Rueter, "The Cybersecurity Dilemma." For a skeptical view, see Liff, "Cyberwar."

28 "Pentagon to Boost Cybersecurity Force," *Washington Post*, January 27, 2013; "Obama Issues Guidance on Cyberwarfare," *Washington Post*, November 15, 2012; "U.S. Military Goes on Cyber Offensive," *Defense News*, March 24, 2012. To be clear, in an offense-dominant cyber world, the pursuit of offensive capabilities would not necessarily signify aggressive intent; rather, doing so would appear to make good strategic sense, given the relative cost effectiveness of offense versus defense. Moreover, if an aggressive adversary has a weak ability to fend off cyber attacks, building the means to target that weakness could bolster deterrence.

29 Lord and Sharp, *America's Cyber Future*, pp. 29.

30 Lord and Sharp, *America's Cyber Future*, pp. 28.

31 "Panetta Warns of Dire Threat of Cyberattack on U.S.," *New York Times*, October 11, 2012.

32 Andrew F. Krepinevich Jr., "The Pentagon's Wasting Assets," *Foreign Affairs* (July/August 2009); Richard Clarke and Robert Knake, *Cyber War: The Next Threat to National Security and What to Do About It* (New York: Harper Collins, 2010).

33 Keir A. Lieber, "The New History of World War I and What It Means for International Relations Theory," *International Security*, Vol. 32, No. 2 (Fall 2007), pp. 155–191; Keir A. Lieber, *War and the Engineers: The Primacy of Politics over Technology* (Ithaca, N.Y.: Cornell University Press, 2005); and Keir A. Lieber, "Grasping the Technological Peace: The Offense-Defense Balance and International Security," *International Security*, Vol. 25, No. 1 (Summer 2000), pp. 71–104.

34 Thomas G. Mahnken, "Cyber War and Cyber Warfare," in Lord and Sharp, *America's Cyber Future*, pp. 57–63, at p. 62.

35 See Dan Reiter, "Exploding the Powder Keg Myth: Preemptive Wars Almost Never Happen," *International Security*, Vol. 20, No. 2 (Fall 1995), pp. 5–34.

36 For a discussion of the danger of preventive war, see Thomas C. Schelling, *Arms and Influence* (New Haven, Conn.: Yale University Press, 1966), pp. 221–259; and Jack S. Levy, "Declining Power and the Preventive Motivation for War," *World Politics*, Vol. 40, No. 1 (October 1987), pp. 82–107.

A Repertory of Cyber Analogies

Robert Axelrod
Ford School of Public Policy, University of Michigan

This report provides a repertory of 35 analogies relevant to issues related to cyber conflict. Analogies such as these can serve several purposes: to motivate (by fear or inspiration), to demonstrate what is possible, to provide examples from the past of things to avoid, and to illuminate particular features of past events that might be worth thinking about in preparation for cyber conflict. The report provides the implications of each analogy. These implications can be thought of as lessons from the past that can be useful once again, despite important changes in technology, doctrine, organization and political context. The analogies are organized in sections covering historical analogies from before, during, and after World War II, and include a section on functional analogies such as those inspired by biology. The report includes an appendix on a tactic that has been used by the Chinese that is quite distinct from Western conceptions of deterrence, namely the denial of retaliatory intent.

BEFORE WORLD WAR II

1. David and Goliath

Asymmetric cyber warfare can topple a giant. The religious significance of the story is somewhat different, namely that overwhelming odds can be overcome if God is on one's side. For many Muslims, the Soviet defeat in Afghanistan is a clear example of David and Goliath.

2. "Remember the Maine"

Destruction of a military target can lead to jingoism and be used to justify war, even if the destruction could have been an accident. In 1898, the *U.S.S. Maine* exploded in Havana harbor, leading to the battle cry of "Remember the Maine, to Hell with Spain!" The exploitation of this event in the American press helped launch the Spanish-American War. The lesson for the cyber realm is that the meaning of an ambiguous event can be shaped by the media to appear to be a deliberately provocative act, resulting in a demand for an overwhelming response. In the case of China, the national media is controlled by the regime, but when the indignation of ultra-nationalist micro bloggers resonates with the broader public, the resulting pressure on the government could be intense. (See also 16, "Gulf of Tonkin Incidents.")

3. Demise of Piracy

Major powers working together can eliminate private attempts to do damage to the global economy, e.g. by holding companies hostage. On the other hand, territories without proper governance (whether geographic or virtual) can be havens for piracy.

4. Privateering

Implication: Some activities that look like piracy might be legally sponsored by a nation, as specified in the U.S. Constitution. "Letters of marque and reprisal could be the latter-day equivalent to empower cyber-privateers in this way to go after certain targets … The possibility of cyber militias comes to mind as well (the Chinese are actually encouraging the formation of these)."[1]

5. Unrestricted Submarine Warfare

A tactic that begins by being regarded as "sneaky" and dishonorable can become accepted, as unrestricted submarine warfare did during the course of World War I. Observing the requirement to warn ships about to be attacked greatly reduced the value of the submarine, which is why giving warning was abandoned.[2] Of course, on the way to becoming accepted, the dishonorable tactic can contribute to an overwhelmingly hostile response. For example the German declaration of

unrestricted submarine warfare contributed to the American decision to enter World War I. The lesson for cyber conflict is that new modes of attack are often seen as dishonorable, and therefore could elicit stronger responses than would otherwise be expected.

6. Unanticipated Information Requirements for British Economic Warfare in World War I

The implementation of a policy may require information in ways not anticipated in peacetime. At the outset of World War I, when Britain tried to implement economic warfare against Germany, it found that the information collected in peacetime was not always what was needed for wartime. In addition, even when the information was collected, it was often distributed over many different parts of the government (and the private sector) in ways that made it impossible to aggregate in a timely way. The difficulties of aggregation included incompatibilities of definitions, periods covered, and formats in which the data is kept. There was the additional problem of withholding information for competitive reasons (either profit or bureaucratic power), as well as legal constraints on sharing.[3]

7. Collateral Damage in British Economic Warfare in World War I

Collateral damage may require restrictions in the use of an otherwise successful form of warfare. Prior to World War I, the British Admiralty had plans to exploit Britain's dominant position in global trade and finance to strangle Germany and its allies at the outbreak of war. When war came, the policy was implemented, but it turned out to be impossible to strangle Germany without impinging on neutral rights in a manner highly provocative to the United States.[4] U.S. cyber measures could hurt allies and neutrals in a conflict (e.g. Japan or EU) so much that the U.S. would have to call off its attack, just as the British had to. An example might be if a conflict that included cyber attacks left only limited bandwidth that was fully secure (say because it went through the latest generation of communications satellites). Then one could imagine that the Pentagon would want to commandeer virtually all of it, but our own private sector and our allies would demand some for themselves.

DURING WORLD WAR II

8. Blitzkrieg

New doctrine is as important as new technology. The French in 1940 had tanks, airplanes and radios, but only the Germans had the doctrine to take advantage of them.

9. Battle of Britain

A conflict can take place entirely within a single domain, such as air-to-air combat or cyberspace. The analogous conflict in cyberspace would be a standalone, overt cyber battle or war between nations, fought entirely within the domain of cyberspace and fully engaging each side's cyber attackers and defenders (probably both in government and the private sector). Though tactical engagements might take place "at the speed of light" these would be mere dogfights in the context of the larger fight, with complete operations as part of offensive and defensive campaigns. A cyber Battle of Britain may develop slowly, through various phases (as did the original, 70 years ago) moving up from smaller, less-organized attacks before blossoming into a full force-on-force unleashing of violence. Each side may be deterred from making larger cyber attacks (as the Germans originally refraining from attacking cities) but continue to one-up the other nation in a progression of violence.[5]

10. Fort Eben Emael

When design criteria are specified too narrowly a supposedly well-designed defense can be easily overcome. On May 10, 1940, seventy-seven Germans in gliders descended on Belgium's strongest fort, Eban Emael. Within a day, they had taken decisive steps to capture the garrison of 1,200. The fort was well "buttoned up" and protected by massive casements and embrasures, being optimized against heavy attack from a distance. To the Belgians, worried about such an attack, gliders were an unknown unknown. Had the Belgians considered an attack from gliders directly onto the fort as even a very unlikely possibility, they could have easily taken measures to defeat a few dozen fully exposed soldiers.[6]

11. Die Glückliche Zeit (Golden Time)

At the start of a major conflict one side might present numerous easy targets until it adapts. *Die Glückliche Zeit* (Golden Time) is the German term for the period in the first summer of World War II when their submarines were able to sink 282 Allied ships.[7] The second Golden Time was the summer after the U.S. entered WWII when German subs were able to sink 609 ships totaling more than three millions tons, roughly a quarter of the tonnage they sunk in the entire war.[8] This is an example of

> the 'harbor lights' phenomenon: when the U.S. entered WWII, it kept eastern seaboard cities' lights on after dark, illuminating targets for U-boats. The lights stayed on for fear of the economic consequences of blackout——and blackout was only imposed when U-boat depredations became too costly. A bit like the cybersecurity problem today. The harbor lights are on all over cyberspace, but the hacker/U-boat captains haven't done enough damage yet to cause more serious security measures to be taken.[9]

12. Raid on Taranto

Vulnerabilities in one's defenses can be revealed by observing how a similar defense was overcome in another setting. On November 11, 1940, British torpedo planes overcame the defenses of the Italian battleships at Taranto by adapting their torpedoes to be effective in shallow waters.[10] Nevertheless, the American Navy failed to learn from this surrogate experience, so the battleships at Pearl Harbor remained vulnerable to Japanese torpedo plane attack.

13. Pearl Harbor

The trauma of Pearl Harbor means that the U.S. will always be alert to the possibility of a "bolt from the blue," even though Pearl Harbor itself was hardly an example of one. In fact, an important lesson of Pearl Harbor (and many other surprise attacks) is that a country can be surprised by the nature of the attack, but is almost never attacked without days, if not weeks, of a serious political crisis that makes war a real possibility in the near future.[11] The implication for cyberwar is that even though a cyber attack can be launched without tactical warning, it is very likely that any major attack will only happen in the context of a serious political crisis. An important lesson is that a potential target of a major cyber attack should be prepared to take advantage of the time available in a crisis to upgrade defensive capabilities in ways that may not be practical in ordinary times. Of course, there may be political constraints on taking any measures that could be seen by the other side as preparations for a preemptive attack.[12]

AFTER WORLD WAR II

14. China Crosses the Yalu

Sometimes attacks are made at the start of a conflict, and then quickly stop. The natural interpretation by the target is that the attack was halted when its initial efforts were thwarted by effective defenses. However, another possibility is that the attacker planned all along that the attack would be launched and then halted in order to send a warning. The attack might have been meant as a demonstration of willingness to resort to this type of attack, and the pause might be designed to give the other side a last chance to avoid a more serious conflict.

In the autumn of 1950 as the United States forces were routing the North Korean Army and racing toward the Yalu River on the border with China, the Chinese tried to warn by both public and private messages that approaching the border would not be tolerated. The Chinese first sent troops over the Yalu to make contact with U.S. forces, and then deliberately broke contact. The U.S. did notice this but did not see it as a warning, despite other numerous diplomatic and public attempts by the Chinese to warn the U.S. that it was about to intervene unless the U.S. backed off.[13]

The Chinese did the same thing against India in 1962, and against Vietnam in 1979.[14] In all three cases, the Chinese warned, then struck in a restrained manner, then paused, and—when their warnings were not heeded—they attacked in strength.[15] I know of no other country that has used this tactic.

The lesson is that when a cyber attack is halted, there are three possible interpretations: the attack failed, the attack was meant as a warning but was actually a bluff, or the attack was meant

as a warning and was not a bluff. Especially if the attack comes from China, the third possibility needs to be taken seriously. [See also the Appendix.]

15. Vietnam and the Tet Offensive

A cyber attack could take the form of guerrilla warfare involving a few large-scale incidents with large-scale effects, but also a continuing string of attritional attacks seeking to erode an adversary's power, influence, and will. A typical tactic of guerrillas is to cause an overreaction from the other, more powerful, adversary as this can help push more people to support the guerrillas' cause. Another is to ensure civilians are impacted directly or indirectly to force them to pressure their government to cease hostilities or influence the way the war is fought. In a true "cyber Vietnam" the attacking group would also have the backing of a national sponsor, aiding and encouraging its campaigns, though possibly unwilling to commit their own cyber or traditional military forces.[16] The massive Tet Offensive of 1968 was based on the premise that the urban population would rise up against the Saigon government if given the chance. The premise turned out to be wrong and the immediate result was the decimation of the Viet Cong. But the Tet Offensive had the unforeseen and possibly decisive effect of undermining the will of the American public to prosecute the war. The lesson for cyber attacks is that the effects may be important without being foreseen.

16. The Gulf of Tonkin Incidents

Seemingly solid information that an attack happened might have been subject to bureaucratic processes that filtered out contradictory information, leading to the same result as if the attack had happened. The first incident in the Gulf of Tonkin was an attack on a U.S. destroyer by ships from North Vietnam on August 2, 1964. Two days later another attack was reported, but that report was based on misinterpretation of radar imagery—an error that was quickly identified but not corrected until much later.[17] In any case, only three days after the second "incident" Congress passed the Gulf of Tonkin Resolution that provided justification for presidential action for the rest of the Vietnam War. The lesson for the cyber realm is that evidence of an attack needs to be verified with care.

17. The Cold War

The lesson most Americans have derived from the Cold War is that a patient policy of containment was successful. The implication of invoking the Cold War is that rivalry can be limited and crises need not explode.[18]

18. Mutual Assured Destruction

Deterrence of nuclear war, and even direct combat between the superpowers, was (apparently) effective for sixty years, making deterrence a highly salient concept for the prevention of cyber war. Despite the attempts to adapt the concept of deterrence to the differences between kinetic and cyber conflict, it has been a stretch. For example, the core concept of mutual assured destruction does not apply. Likewise, the core concept of deterrence that requires clarity of response in order to achieve credibility of commitment does not necessarily apply to cyber conflict, since ambiguity might be helpful to avoid retaliation, even if the ambiguity lessens deterrence. (For more on China's use of ambiguity, see analogy 24, "Chinese Restriction of Rare Earth Exports," and the Appendix.)

19. Escalation Ladder

The clarity of the nuclear threshold has helped sustain the taboo against the use of nuclear weapons. There are potential thresholds between cyber espionage and cyberwar, but they are not yet widely understood or agreed upon. Nor is there even convergence on how the terms should be defined. There is not even convergence on what kinetic actions in response to a given kind of cyber attack would constitute escalation or de-escalation.

20. Control of Chemical Weapons

Even without effective verification, agreements on limiting cyber attacks (e.g. to military targets) could prove effective.

21. MIRV (multiple independently targetable re-entry vehicle)

In a rivalry, the side with a technical advantage (such as the U.S. had in the 1970s with MIRVs) may miss an opportunity to prohibit a destabilizing technology. In the case of MIRVs,

detection was easy at the stage of testing, but almost impossible once deployed. In retrospect, the U.S. would have been better off with an early arms control treaty banning the testing of this destabilizing technology. At the time, however, the well-established principle prevailed that a military advantage should never be voluntarily surrendered, and there was little or no consideration of the destabilizing potential of that technology. The analogy will be apt when one side has a lead in a technology that would be destabilizing if deployed by both sides of a rivalry, when prohibition would be more reliably verifiable before deployment than afterwards.

22. 9/11

There are terrorists who are plotting to inflict maximum damage to the U.S. population. Beside the obvious possibility of nuclear or biological weapons, there is also the potential danger of a cyber attack on a critical target such as a dam, hospital, power grid, or water purification system.

23. Wikileaks

When classified information is widely distributed to promote the "connection of the dots", there is a corresponding risk of massive leakage.

24. Chinese Restriction of Rare Earth Exports

An act by one country that harms another is often ambiguous in its intent, even if the effect and the perpetrator are both clear. On September 7, 2010 the Japanese detained the captain of a ship in waters around a disputed island. The Chinese then cut their exports of rare earths by 72%. Since rare earths are essential for a variety of electronic and other industrial products, and China controlled 95% the global supply, the timing of the export restriction was seen by many in Japan and the West as retaliation, despite Chinese denials.[19] For more on the Chinese tactic of denying retaliatory intent, see the Appendix.

25. Cyber Espionage

Espionage is done by everyone and is not an act of war. Nations maintain a "polite fiction" that they don't do it, even if their rivals do. The burden is on the defense. Revealing espionage often harms bilateral relations. The amount of harm done by cyber espionage, especially by China, is substantial but the U.S. public has not been aroused.

26. Cyber Attack on Siberian Pipeline

Cyber industrial sabotage by means of malware is nothing new. In 1982, the CIA introduced a logic bomb into exported pipeline software that was picked up by the KGB, leading to "the most monumental non-nuclear explosion and fire ever seen from space."[20]

27. Cyber Attacks on the Iranian Nuclear Program

The use of cyber attacks by the U.S. and Israel against infrastructure (as opposed to cyber espionage) now has a precedent,[21] making it easier for other nations to justify another such attack. A potential "red line" still exists for attack on financial systems.

28. DigiNotar Certificate Authority Breach

Even the most trusted category of cyber authority could have "shocking ineptness" in its security system. DigiNotar was a supplier of trusted certificates to authenticate that a request on the Internet was being sent to the intended party. In 2011, over 500 false certificates for domains such as Google and Yahoo were issued through DigiNotar by an Iranian hacker. This hack resulted in 600,000 requests that were subject to a "man-in-the middle" attack. Over 95% of these requests came from Iran, suggesting that the purpose was to spy on Iranian Internet users. After the fact, an audit showed that DigiNotar's

> servers ran out-of-date software. Its network was poorly segmented, so problems would not be contained if they arose. Passwords in play at the time of the hack might easily have been guessed via brute-force attack. In addition, there was no secure logging and server-side antivirus protection was absent.[22]

In my opinion, some or all of these failures in elementary security practices would have been known and must have been

tolerated by co-workers. The most important lesson is that cybersecurity indoctrination should include a version of West Point's Honor Code such as "I will not violate cybersecurity procedures, or tolerate those who do."[23]

FUNCTIONAL ANALOGIES

29. Biodiversity vs. Weakest Link

The biodiversity metaphor suggests that diversity of cyber systems may result in resilience against attacks. On the other hand, if the problem is to protect information stored in various systems, the "weakest link" metaphor suggests that diversity of cyber systems makes the defense as weak as its weakest component.

30. Herd Immunity

If a sufficient proportion of the population is immune to a disease, the disease is unable to spread among the vulnerable parts of the population.

31. Crime

A wide range of private (as opposed to state-sponsored) cyber activity can be suppressed by ordinary police work and the criminal justice system.

32. Child Pornography

Some things are universally abhorred, and such things could be the basis of initial understandings and norms about activities in cyberspace.

33. Territorial Responsibility

The legal principle that a state is responsible for the prevention of illegal acts emanating from its territory can be extended to cyberspace to hold nations responsible for cyber activity launched from their own territories. While the origin of a cyber activity is often impossible to trace, there may be times when its origin *can* be established.

34. World Trade Organization

The principle of equivalent retaliation built into the treaty of the World Trade Organization makes enforcement of its rulings quite effective. If such a principle could be established for violations of norms in the cyber world, self-help enforcement could also be effective.

35. Insurance and Industry Standards

Individuals and companies purchase insurance to mitigate the effects of theft and other crimes. In turn, insurance companies often set standards that require certain anti-theft measures to reduce their liability. Insurance against cyber crime is not a well-established industry, largely because of the difficulty of assessing damage from a cyber crime in monetary terms. Nevertheless, there may be value is exploring whether the standards required by insurance companies could be adapted to prevent cyber crime. For example, the computer security industry could set standards and issue the equivalent to a "Good Housekeeping Seal of Approval" for companies that meet those standards.

APPENDIX

Chinese Tactic of Denying Retaliation

The historical analogy Number 24, "Chinese Restriction of Rare Earths Exports" is worthy of elaboration because it involves an unusual tactic that China has used several times recently, and is readily adaptable to cyber conflict.

In the last few years, China has employed a new pressure tactic against three countries with which it has a dispute: Japan, North Korea, and the Philippines. In each case, China suspended trade in specific commodities, while refusing to acknowledge that the trade suspension had anything to do with the dispute. In two of the cases, China has apparently achieved its immediate goals, and the third case is still unfolding.

- After a Chinese ship captain was detained in Japan for sailing in waters near a disputed island on September 7, 2010, China drastically curtailed its exports of rare earths. Rare earths are important in the manufacture

of many electronic products, and China controlled 95% of the global supply.[24] China denied it had a trade embargo with Japan, but after the captain was released, the trade returned to normal.[25]

- In January 2011, China suspended oil supplies to North Korea following the North's shelling of Yeonpyeong Island. This was widely interpreted as an effort to prevent Pyongyang from carrying out its threats to retaliate against the South if the South went ahead with its live-fire exercises as planned.[26] China has not publicly acknowledged its oil cut off, let alone provided a reason. Earlier suspensions without public acknowledgement have apparently occurred in 2003[27], 2006[28], and 2008[29]. For example, in March 2003, China suspended oil shipments to North Korea for three days due to "technical difficulties" soon after Pyongyang test-fired a missile into waters between Korea and Japan. The move was widely interpreted as a successful effort to get North Korea to attend a trilateral meeting in Beijing the following month.
- On April 10, 2012, a Philippine naval ship tried to arrest Chinese fishermen near a disputed reef in the South China Sea. China then refused to allow 150 containers of bananas to enter its market, saying that the bananas were "crawling with insects." The Philippines denied the charges and said that the insects the Chinese cited attack coconuts, not bananas.[30] China never acknowledged that its interruption of trade with the Philippines was linked to the territorial dispute.

Four questions arise with respect to these cases: What's new in the Chinese tactic? Why deny? Why China? and What's next?

What's New?

Countries have frequently resorted to economic pressure to get their way on some dispute. What is new in the Chinese tactic is the refusal to acknowledge that the pressure has any relationship to the issue at hand. I can think of no other country using a trade disruption to provide pressure on a security issue, where the timing of the disruption was publicly presented as totally coincidental.

Of course, other countries have often used economic pressure to attain security goals. For example, in 1956 when Britain and France invaded Suez, the United States successfully used financial pressure to force them to withdraw. But the United States did not claim that its financial sanctions were merely coincidental. Nor has Pakistan claimed any pretext when it expressed its anger at U.S. actions by halting NATO supply trucks en route to Afghanistan in 2010 and again in 2011.[31]

There are also many cases in which a country took military action that it did not acknowledge, or sought "plausible deniability." The U.S. responsibility for the Bay of Pigs invasion is just one of many examples, some successful and some not.[32] But I can't think of any incidents in which the actions in the economic domain were taken to apply pressure in the security domain, along with claims that the timing of the economic pressure was purely coincidental.

In fact, standard strategic doctrine—as understood in the West—emphasizes that threats and warnings should be explicit for two reasons: to achieve maximum credibility, and to make clear what must be done to end the pressure. This raises the questions of why one might deliberately deny that a trade disruption is related to the security issue at hand, and why China is the one using this new tactic.

Why Deny?

Apparently the purpose of denying that the trade disruption is related to the security issue is to allow the other side to save face when backing down. Even if everyone knows that there is a linkage, the idea that there isn't any linkage is something we might call "a polite fiction."[33]

Polite fictions are common in everyday discourse such as the polite fiction "All teachers at our school admire one another and the principal." Everyone knows or suspects this is a fiction, but the statement's veracity is never pressed. It serves like the willing suspension of disbelief—allowing everyone to maintain the personae they have constructed for the purpose of social interaction.[34]

In blunt strategic terms, the polite fiction of the Chinese tactic of denying that undue pressure is being brought to bear lowers

the cost to the other side of backing down—something of obvious value to the Chinese.[35]

Why China?

It is often said that East Asian cultures are more concerned with "saving face" than Western cultures are. Perhaps so, but there are plenty of examples in which Western countries have put great store in saving face.[36] For example, in the Cuban Missile Crisis President Kennedy took care to call his action a "quarantine" rather than a "blockade" because a blockade was an act of war and he did not want the Soviets to have to acknowledge giving in to an act of war. Even more important, in the deal that resolved the crisis, the Americans insisted to the Soviets that the promised removal of American missiles from Turkey remain secret, so that neither the U.S. nor its Turkish ally would lose face when the missiles were actually removed a few months later.[37]

So if other countries have also been concerned with saving face, why has China been the one to invent the tactic of claiming that the timing of its economic pressure was only coincidentally related to a security issue? One reason is that China is concerned to support its claim that it seeks a "peaceful rise". For this reason it wants to avoid acknowledging that it uses undue pressure to resolve security issues. Another reason why China, rather than a Western power, is the one to invent this tactic is that (as described earlier), ambiguous threats and warnings are simply inconsistent with the dominant Western conception of how to achieve deterrence and compellance. One might want to be a bit vague about the consequences if things escalate, but one wouldn't want to leave any unnecessary doubt in the target's mind that a threat was being issued, and one would want to display as much commitment as possible that further action would be taken if the situation remained unsatisfactory. Or so says standard Western security doctrine.

Indeed the Western approach to clarity draws not only on game theory, but also on major lessons from the outbreak of the two most traumatic events in the West, namely World War I and World War II. At the outbreak of World War I, Britain had not yet made clear that it would declare war on Germany if Germany violated the neutrality of Belgium. An important lesson was that clarity might have deterred Germany from invading Belgium.[38] Likewise, a major lesson from the failure to deter Germany from launching World War II is that the Allies should have decided much earlier and made it very clear that they would resist Hitler's aggressive demands by force if necessary. On the other hand, China's experience—both before and after 1949—is that subtlety is often better than clarity.

What Next?

China's use of its new tactic has clearly achieved its immediate goal when applied to both Japan and North Korea, but it is too early to tell if it achieved its immediate goal when applied to the Philippines. But it is plausible to assume that the tactic works well at a low enough cost to China and would be used again when the conditions are right. The conditions seem to be that China wishes to exert pressure in a given domain (such as a security issue), but wants to avoid the appearance of using pressure. The desire to avoid the appearance (or at least the acknowledgement) of pressure can be due to several factors, including China's desire to maintain its posture of "peaceful rise," its desire to avoid domestic reactions from its own public or the publics of the targeted country, and its desire to make it easier for the other side to give in to China. No doubt these conditions are likely to arise many times in the years to come, not only on issues related to sovereignty over disputed islands, but on other issues of deep concern to China in dealing with countries like North Korea, and perhaps Taiwan.

China must, however, weigh the prospects of short-term success against the possibility of incurring long-term costs with its new tactic. For example, China's disruption of rare earth exports was quickly followed by Japan's release of the Chinese sea captain it held, but it also led to a global awareness of China's virtual monopoly of the supply of these valuable materials.[39] The result has been a buildup of inventories of rare earths and a readiness to restore production elsewhere, two steps that will soon dramatically reduce the vulnerability of other countries to any future disruption of Chinese exports of rare earths. In retrospect, China may regret not having saved its one-time opportunity to exert this pressure in a dispute of greater importance to China.[40] They may also come to regret having escalated pressure on the Philippines, the result

of which may be greater U.S.-Philippines security cooperation—albeit under the polite fiction that it has nothing to do with China.[41]

In a future confrontation with the United States, a country might choose to use a cyber attack rather than an economic action. A cyber attack could be designed both to show displeasure with the United States, and to imply the possibility of escalation if it is not satisfied with the American response. A cyber attack has the advantage of not being as easily attributable as an economic action would be. To make it easier for the U.S. to give in, the instigator may once again assert that whatever harm occurred was not intended, and that the timing was purely coincidental.[42] ❇

NOTES

1 John Arquilla, personal communication.

2 Nicholas Lambert, personal communication.

3 Nicholas A. Lambert. *Planning Armageddon: British Economics Warfare and the First World War*. (Cambridge: Harvard University Press, 2012)

4 Ibid.

5 Adapted from Gregory Rattray and Jason Healey, "Categorizing and Understanding Offensive Cyber Capabilities and Their Use" in Proceedings of a Workshop on Deterring CyberAttacks, Committee on Deterring Cyberattacks, National Research Council. 2010. http://sites.nationalacademies.org/xpedio/groups/cstbsite/documents/webpage/cstb_059437.pdf

6 See references in the Wikipedia article on "Fort Eben-Emael".

7 See http://ww2db.com/battle_spec.php?battle_id=277

8 See sources at http://en.wikipedia.org/wiki/Second_Happy_Time

9 John Arquilla personal communication

10 Angelo N. Caravaggio, "The Attack at Taranto," *Naval War College Review*, vol. 59, no. 3. (2006) Pp. 103–127.

11 Other well-known surprises support this point. For example, the U.S. was surprised by China's entry into the Korean War, but see item 14. Stalin was surprised by Hitler's attack, but that was because Stalin discounted the extensive evidence he had that an attack was imminent. Israel was surprised in 1973, but it too had sufficient warning that its leaders chose to discount. In all these cases, the attacks were preceded by days, if not weeks, of a serious political crisis that made war a real possibility in the near future.

12 Alternatively, conspicuous preparations for escalation can sometimes help *deter* the other side from pursuing the conflict. For example in the Cuban Missile Crisis, the conspicuous preparation the U.S. undertook to invade Cuba was one of the main reasons why Khrushchev decided to end the crisis by withdrawing the missiles.

13 Allen Whiting, *China Crosses the Yalu* (New York: Macmillan, 1960).

14 Yee, Herbert S. "The Sino-Vietnamese Border War: China's Motives, Calculations and Strategies." *China Report* (Jan.-Feb. 1980), Pp. 15–32.

15 Allen S. Whiting, 2001, "China's Use of Force, 1950–96, and Taiwan," *International Security*, Vol. 26, No. 2 (Autumn, 2001), pp. 103–131.

16 Adapted Gregory Rattray and Jason Healey, "Categorizing and Understanding Offensive Cyber Capabilities and Their Use" in Proceedings of a Workshop on Deterring CyberAttacks, Committee on Deterring Cyberattacks, National Research Council. 2010. http://sites.nationalacademies.org/xpedio/groups/cstbsite/documents/webpage/cstb_059437.pdf

17 Robert Hanyok. "Skunks, Bogies, Silent Hounds, and the Flying Fish: The Gulf of Tonkin Mystery, 2–4 August 1964, *Cryptologic Quarterly* (1998). The declassified version of this report by the NSA Historian is available at http://www.gwu.edu/~nsarchiv/NSAEBB/NSAEBB132/relea00012.pdf

18 For other analogies from the Cold War, see David Sulek and Ned Moran, "What Analogies Can Tell Us About the Future of Cybersecurity." Pp. 118–131 in *The Virtual Battlefield: Perspectives on Cyber Warfare*, vol. 3, *Cryptology and Information Security Series*, edited by C. Czosseck and K. Geers. (Amsterdam: IOS Press, 2009).

19 "China Denies Japan Rare-Earth Ban Amid Diplomatic Row," *Bloomberg News*, September 23, 2010.

20 Thomas C. Reed, "At the Abyss: An Insider's History of the Cold War," (New York: Ballantine Books, 2004). See also sources cited in the Wikipedia article on "Siberian pipeline sabotage".

21 David Sanger, "Obama Order Sped Up Wave of Cyberattacks Against Iran," *New York Times*, June 1, 2012.

22 John Leyden, 2011. "Inside 'Operation Black Tulip': DigiNotar hack analysed" *The Register*. http://www.theregister.co.uk/2011/09/06/diginotar_audit_damning_fail/
See also Fox-IT (August 2012). *Black Tulip: Report of the investigation into the DigiNotar Certificate Authority breach.*

23 What is needed is a metanorm, i.e., a norm against tolerance of norm violations. See Robert Axelrod, 1986. "An Evolutionary Approach to Norms," *American Political Science Review*, vol. 80, pp. 1095–1111. http://www-personal.umich.edu/~axe/Axelrod%20Norms%20APSR%201986%20(2).pdf

24 http://www.bloomberg.com/news/2010-09-23/china-denies-japan-rare-earth-ban-amid-diplomatic-row-update1-.html

25 http://online.wsj.com/article/SB10001424052702303879604577409831672468236.html

26 http://www.koreatimes.co.kr/www/news/nation/2011/01/117_79966.html. Earlier suspensions of oil shipments to North Korea apparently took place in 2006 and 2008. http://www.nytimes.com/2006/10/30/world/asia/30iht-oil.3334398.html and nautilus.org/napsnet/napsnet-special-reports/dprk-prc-trade-aden/

27 http://www.armscontrol.org/act/2009_07-08/zhang

28 http://www.nytimes.com/2006/10/30/world/asia/30iht-oil.3334398.html?_r=1

29 http://nautilus.org/napsnet/napsnet-special-reports/dprk-prc-trade-aden/

30 http://www.csmonitor.com/World/terrorism-security/2012/0515/Philippines-feels-the-economic-cost-of-standing-up-to-China

31 http://www.juancole.com/2010/10/pakistan-opens-khyber-crossing-to-nato-supply-trucks-but-issues-threats-over-hot-pursuit.html and http://articles.cnn.com/2011-11-27/asia/world_asia_pakistan-nato-attack_1_nato-helicopters-khyber-agency-nato-trucks?_s=PM:ASIA

32 The Chinese certainly believe that the U.S. attack on their embassy in Belgrade on May 7, 1999 was an example of a deliberate attack that was presented to the world as a mistake. For evidence that it was deliberate see John Sweeney et al. "Nato bombed Chinese deliberately," *The Guardian/The Observer,* Oct. 16, 1999. http://www.guardian.co.uk/world/1999/oct/17/balkans

33 In the context of international relations, the concept of a "polite fiction" was apparently first used to describe the obviously false claim by the Soviets that they never engaged in spying. Robert Axelrod and William Zimmerman, "The Soviet Press on Soviet Foreign Policy: A Usually Reliable Source," *British Journal of Political Science, 11* (April 1981), pp. 183-200.

34 http://www.justmusing.net/2010/01/26/polite-fiction/

35 For a formal game theoretic model in which "saving face" is important, see Barry O'Neill, *Honor, Symbols and War*, 1999 (Ann Arbor, MI: University of Michigan Press).

36 For many examples, see Barry O'Neill, *Honor, Symbols and War*, (Ann Arbor, MI: University of Michigan Press, 1999), especially pp. 139–63.

37 Israel also provides examples of not acknowledging its actions, even when there is no pretense of plausible deniability. For example, Israel has not acknowledged its possession of nuclear weapons, or its 2007 aerial attack on a Syrian nuclear facility. In both cases, an important goal is to allow other parties to avoid having to respond to Israel's actions.

38 See for example Barbara Tuchman, *The Guns of August* (New York: Macmillan, 1962)

39 An interesting comparison with the Sino-Japanese territorial dispute is the Korean-Japanese territorial dispute. The latter has involved various forms of pressure but all have been directly related to the dispute rather than indirect or unacknowledged pressure in some other domain such as trade. See for example the sources in Wikipedia's "Liancourt Rocks dispute."

40 On when to use a potentially fleeting cyber resource, see Robert Axelrod and Rumen Iliev (2014), "The Timing of Cyber Conflict," *Proceedings of the National Academy of Sciences*, Vol. 111.

41 For example, earlier pressure from China let Secretary of State Hillary Clinton to say, "our long mutual defense treaty and alliance relationship with the Philippines [requires] working with the Philippines to provide greater support for external defense particularly maritime domain awareness, defensive ones, maritime boundaries." Williard Cheng, "Clinton Heaps Praise on Pacquiao, reaffirms U.S. support for PH," *ABS-CBN News*, Nov. 11, 2011.

42 This work was supported in part by Air Force Office of Scientific Research Grant FA9550-10-1-0373.

Editors' Acknowledgments

The editors are deeply indebted to all of the contributors to this Technical Report. Their many insights remind us of the power of analogical thinking, and are sure to inform and guide the ongoing discourse about cyber strategy. Special thanks are owed to General Keith Alexander. His vision, encouragement, and support were essential to the design, progress, and completion of this undertaking.

At the Naval Postgraduate School, thanks are owed to Dr. Hy Rothstein, Director of the Department of Defense Information Operations Center for Research, for embracing the idea of publishing this anthology. Rebecca Lorentz managed the whole process, taking on a wide range of tasks with skill and grace. Design and layout are by Ryan Stuart; Amelia Simunek and Major Simon Powelson also played key roles in producing artwork.

About the Contributors

Robert Axelrod is the Walgreen Professor for the Study of Human Understanding at the University of Michigan. Dr. Axelrod is a member of the National Academy of Sciences as well as the Council on Foreign Relations. His books *The Evolution of Cooperation* (Basic Books, 1984) and *The Complexity of Cooperation: Agent-Based Models of Competition and Collaboration* (Princeton, 1997) have been translated into a dozen languages. In 2013 Dr. Axelrod received the Lifetime Achievement Award from the American Political Science Association's Conflict Processes Section for "a body of work unmatched within conflict processes in terms of its breadth, influence, and creativity."

Dorothy Denning is Distinguished Professor of Defense Analysis at the U.S. Naval Postgraduate School. Previously she served as Director of the Georgetown Institute of Information Assurance. She has testified before the U.S. Congress on encryption policy and cyber terrorism, and has served in leadership and advisory positions with government agencies and private sector organizations. Her books include *Information Warfare and Security* (Addison-Wesley, 1998), and *Internet Besieged: Countering Cyberspace Scofflaws* (ACM, 1997). Dr. Denning was recently inducted into the International Cyber Security Hall of Fame.

Peter Feaver is Professor of Political Science and Public Policy at Duke University. He is Director of the Triangle Institute for Security Studies (TISS) and also Director of the Duke Program in American Grand Strategy (AGS). Dr. Feaver is author of *Armed Servants: Agency, Oversight, and Civil-Military Relations* (Harvard, 2003) and of *Guarding the Guardians: Civilian Control of Nuclear Weapons in the United States* (Cornell, 1992). He is co-author: with Christopher Gelpi and Jason Reifler, of *Paying the Human Costs of War* (Princeton, 2009). In 1993–94, Feaver served as Director for Defense Policy and Arms Control on the National Security Council at the White House, where his responsibilities included the national security strategy review, counter-proliferation policy, regional nuclear arms control, and other defense issues.

Kenneth Geers is a Senior Global Threat Analyst at FireEye. Dr. Geers spent twenty years in the U.S. Government, with lengthy tours at NSA, NCIS, and NATO. Kenneth was the first U.S. Representative to the NATO Cooperative Cyber Defence Centre of Excellence in Estonia. He is the author of *Strategic Cyber Security*, Editor of *The Virtual Battlefield: Perspectives on Cyber Warfare*, Technical Expert for the *Tallinn Manual on the International Law Applicable to Cyber Warfare*, and author of more than twenty articles and chapters on cyber conflict. Follow him on Twitter @KennethGeers.

Michael S. Goodman is Reader in Intelligence and International Affairs in the Department of War Studies, King's College, London. He is author or editor of four previous books, including *The Routledge Companion to Intelligence Studies* (Routledge, 2013), *Spying on the Nuclear Bear* (Stanford, 2007) and Volume I of *The Official History of the Joint Intelligence Committee (Routledge, forthcoming).*

John Kao is chairman of the Institute for Large-Scale Innovation. He has worked in the field of innovation for 30 years as an advisor, practitioner, author and educator. Dr. Kao earned an MD at Yale, an MBA from Harvard and is an accomplished jazz musician. He was dubbed "Mr. Creativity" and a "serial innovator" by *The Economist*. His books include the bestselling *Jamming: The Art and Discipline of Business Creativity* (Harper Business, 1997) and *Innovation Nation, How America is Losing its Innovation Edge, Why It Matters and What We Can Do To Get It Back* (Free Press, 2007).

Nicholas Lambert is Associate Fellow of the Royal United Services Institute, Whitehall, London, and a Visiting Professor of History at the University of Maryland. His first book, *Sir John Fisher's Naval Revolution* (Columbia, SC, 1999), was awarded the 2000 Society of Military History's Distinguished Book Award and the Western Front Association Norman Tomlinson Prize. His most recent book is *Planning Armageddon: British Economic Warfare and the First World War* (Harvard, 2012).

Keir Lieber is Associate Professor in the Edmund A. Walsh School of Foreign Service and the Department of Government at Georgetown University. Dr. Lieber's current research focuses on contemporary grand strategy and U.S. nuclear weapons policy. His articles have appeared in leading scholarly and foreign policy journals such as *International Security* and *Foreign Affairs* and he is the author of *War and the Engineers: The Primacy of Politics over Technology* (Cornell, 2005), and *War, Peace, and International Political Realism:* (Notre Dame, 2009).

Gregory Rattray is the CEO and founding partner of Delta Risk, LLC—a risk management consulting firm. Previously he was Commander of the 23rd Information Operations Squadron responsible for U. S. Air Force information warfare tactics and target development. He is the author of *Strategic Warfare in Cyberspace* (MIT, 2001) and editor of *Arms Control Toward the 21st Century* (Lynne Rienner, 1996). Dr. Rattray also served as the director for cyber security on the National Security Council staff in the White House. He was a key contributor to the President's National Strategy to Secure Cyberspace, and helped initiate the first national cyber security exercise program involving government and the private sector.

Bradley J. Strawser is Assistant Professor of Defense Analysis at the U.S. Naval Postgraduate School and a research associate at the Oxford University Institute for Ethics, Law and Armed Conflict. Dr. Strawser's work has appeared in such journals as *Analysis, Philosophia*, *Journal of Military Ethics*, *Public Affairs Quarterly*, *Journal of Human Rights*, and *Epoché.* His recent book, *Killing By Remote Control: The Ethics of an Unmanned Military* (Oxford, 2013), examines the ethical questions surrounding the employment of drones.

John R. "Buck" Surdu retired from the U.S. Army in 2013. One of his last assignments was on the staff of U.S. Cyber Command. In addition to his long military career, including management of technology projects for the U.S. Army and DARPA, he has published cyber-focused articles such as "Army, Navy, Air Force, and Cyber? Is it Time for a Cyberwarfare Branch of the Military?" Information Assurance Newsletter (2009) and "Deep Green: Commander's Tool for COA's Concept" Computing, Communications and Control Technologies (2008).

Michael Warner serves as the Command Historian for U.S. Cyber Command. He has written and lectured widely on intelligence history, theory, and reform. Dr. Warner teaches at American University's School of International Service and at Johns Hopkins University's Advanced Academic Programs, and is on the board of editors of the journal *Intelligence and National Security*. Recent essays include: "Cybersecurity: A Pre-History," *Intelligence and National Security* (October, 2012); "The Rise of the US Intelligence System," in Loch Johnson, ed., *The Oxford Handbook of National Security Intelligence* (Oxford, 2010); and "Building a Theory of Intelligence Systems," in Greg Treverton & Wilhelm Agrell, eds., *National Intelligence Systems: Current Research and Future Prospects* (Cambridge, 2009). His latest book *The Rise and Fall of Intelligence: An International Security History* is being published by Georgetown University Press.

James J. Wirtz is Dean of the School of International Graduate Studies at the Naval Postgraduate School. He is also editor of the Palgrave Macmillan series, *Initiatives in Strategic Studies: Issues and Policies,* and former chair of the Intelligence Studies Section of the International Studies Association. Dr. Wirtz has authored and co-edited books on intelligence, deterrence, the Vietnam War and military innovation and strategy. His books include: *Intelligence: Windows Into a Hidden World* (Roxberry, 2004); *Encyclopedia of Weapons of Mass Destruction* (ABCClio, 2004); *Nuclear Transformation: The New U.S. Nuclear Doctrine* (Palgrave Macmillan, 2005); and *Globalization and WMD Proliferation* (Routledge, 2007).

About the Editors

Emily O. Goldman is a senior advisor at the U.S. Cyber Command and has held a range of positions in the Department of Defense over the past decade. Dr. Goldman was formerly Professor of Political Science at the University of California, Davis. Her books include: *Sunken Treaties: Naval Arms Control Between the Wars* (Penn State, 2002); *The Diffusion of Military Technology and Ideas* (Stanford, 2003); *National Security in the Information Age* (Routledge, 2004); *Power in Uncertain Times* (Stanford, 2010); and *Information & Revolutions in Military Affairs* (Routledge, 2013).

John Arquilla is Professor of Defense Analysis at the U.S. Naval Postgraduate School. He is best known for having developed the original concepts of cyberwar, netwar, and swarm tactics. Dr. Arquilla's books include: *Networks and Netwars* (RAND, 2001); *Worst Enemy* (Ivan R. Dee, 2008); *Insurgents, Raiders, and Bandits* (Rowman & Littlefield, 2011); and *Afghan Endgames* (Georgetown, 2012).

Image Credits

Cover and title page images: The U.S. Navy battleship *USS California* (BB-44) slowly sinking alongside Ford Island, Pearl Harbor, Hawaii (USA), as a result of bomb and torpedo damage, 7 December 1941. The destroyer *USS Shaw* (DD-373) is burning in the floating dry dock YFD-2 in the left distance. The battleship *USS Nevada* (BB-36) is beached in the left-center distance. U.S. Government work, public domain; via Wikipedia at http://en.wikipedia.org/wiki/File:USS_California_sinking-Pearl_Harbor.jpg

6 Fotolia.com

9 USAF Fighters fly over Kuwaiti oil field 1991 Desert Storm by USAF, public domain; via Wikimedia Commons at http://en.wikipedia.org/wiki/File:USAF_F-16A_F-15C_F-15E_Desert_Storm_edit2.jpg

13 Remember Dec 7th by United States Office of War Information, United States National Archives, public domain; via Wikimedia Commons at http://commons.wikimedia.org/wiki/File%3ARemember_december_7th.jpg

16 9/11 damage by Cyril Attias; licensed by Creative Commons 2.0 Generic at http://www.flickr.com/photos/newyork/6113249955/in/photostream/

21 Cyber warriors by USAF, public domain; via Wikipedia at http://pt.wikipedia.org/wiki/Ficheiro:Monitoring_a_simulated_test_at_Central_Control_Facility_at_Eglin_Air_Force_Base_(080416-F-5297K-101).jpg

28 Isoroku Yamamoto & Saddam Hussein, public domain; via Wikimedia Commons at http://commons.wikimedia.org/wiki/File:Isoroku_Yamamoto.jpg

32 Pearl Harbor, National Archives, public domain; via Wikimedia Commons at http://upload.wikimedia.org/wikipedia/commons/6/6f/Pearl_harbour.png

35 Pershing launch by Warren C. Weaver, Civ. U.S. Army Photograph, public domain; via Wikimedia Commons, http://en.wikipedia.org/wiki/File:Pershing_1_launch_(Feb_16,_1966).png

38 Fotolia.com

40 Fotolia.com

42 Power line towers (cropped), by PNNL—Pacific Northwest National Laboratory, licensed by Creative Commons; via Flickr at http://www.flickr.com/photos/pnnl/7404564340/sizes/o/in/photostream/

45 USB drive, by Achim Raschka, licensed by Creative Commons 3.0; via Wikimedia Commons at http://commons.wikimedia.org/wiki/File:12-11-03_intenso_USB_drive.JPG

51 U.S. Special Operations in Afghanistan by U.S. DoD, public domain; via Wikipedia at http://en.wikipedia.org/wiki/File:Laser_designator-_SOF_in_Afghanistan.jpg

52 GPS Satellite, U.S. Government, public domain; via Wikimedia Commons at http://commons.wikimedia.org/wiki/File%3AGPS-IIR.jpg

55 *Inspire* Magazine, screenshot; via Wikipedia at http://en.wikipedia.org/wiki/File:Inspire_magazine_cover.png

60 WW II air raid by USAF National Museum, public domain; via http://www.nationalmuseum.af.mil/photos/media_search.asp?q=ploesti

65 Patriot missiles are launched to intercept an Iraqi Scud missile over the city of Tel Aviv by Alpert Nathan, GPO, licensed under Creative Commons; via Wikimedia Commons at http://commons.wikimedia.org/wiki/File%3AFlickr_-_Government_Press_Office_(GPO)_-_Patriot_missiles_being_launched_to_intercept_an_Iraqi_Scud_missile.jpg

67 Webcam photo of Russian hacker, Georgia, released by Georgian government into the public domain; at http://dea.gov.ge/uploads/CERT%20DOCS/Cyber%20Espionage.pdf

69 Fotolia.com

70 German propaganda leaflet by Nazi German propaganda ministry PROMI (de:Propagandaministerium), public domain; via Wikimedia Commons at http://commons.wikimedia.org/wiki/File%3ANaziV-WaffenRocketPropagandaleaflet1944june.jpg

72 Iron Dome during "Operation Pillar of Cloud" by Emanuel Yellin, licensed by Creative Commons 3.0; via Wikimedia Commons at http://commons.wikimedia.org/wiki/File%3AIronDome246.jpg

73 Fotolia.com

78 Wireless radio station by H. Graumann and I. Piedade Pó, public domain; via Wikimedia Commons at http://commons.wikimedia.org/wiki/File%3AAFDCM-02-017.jpg

80 Exchange rates by Woodarelaisku (Own work), licensed under Creative Commons 3.0; via Wikimedia Commons at http://commons.wikimedia.org/wiki/File%3ASK_Korea_tour_%E9%A6%96%E7%88%BE_%E6%9C%80%E4%BD%B3%E8%A5%BF%E6%96%B9_%E9%A6%96%E7%88%BE%E8%8A%B1%E5%9C%92%E9%85%92%E5%BA%97_Best_Western_Premier_Seoul_Garden_Hotel_money_Exchange_display_July-2013.JPG

78 Hamilton John Agmondesham Cuffe, 5th Earl of Desart by Walter Stoneman, © National Portrait Gallery, London, used by permission

85 *The Sun* newspaper, August 1, 1914, via Library of Congress at http://chroniclingamerica.loc.gov/lccn/sn83030272/1914-08-01/ed-1/seq-1/

88 Fotolia.com

89 Fotolia.com

93 Fotolia.com

95 Fotolia.com

99 Field Marshal Erwin Rommel in North Africa by Bundesarchiv, Bild 101I-784-0249-04A / Koch / CC-BY-SA; via Wikimedia Commons at http://commons.wikimedia.org/wiki/File:Bundesarchiv_Bild_101I-784-0249-04A,_Nordafrika,_Rommel_im_Befehlsfahrzeug_%22Greif%22.jpg