Some Thoughts on Deterrence in the Cyber Era

R Jervis

Columbia University
New York, New York, U.S.A.
E-mail: rlj1@columbia.edu

Abstract: *Although cyber is an increasingly important instrument of national policy for deterrence and coercion, our understanding of how it functions is still weak and underdeveloped. This article explores some traditional notions of deterrence that must be remembered and reinvigorated, and notes aspects peculiar to cyber that must also be explored. Traditional deterrence makes clear the importance of promises and the difficulty in telling what the other side values and fears, and how it sees the world. Cyber conflict poses special problems in the uncertainties and ambiguities involved in, and related to the lack of shared understandings about what would constitute escalation.*

Keywords: *Cyber Deterrence, Intelligence and Cyber, Escalation*

Introduction

The somewhat awkward title of this article underscores two known but sometimes neglected truths: there is no such thing as cyber deterrence, and it is far from clear that cyber should be considered a 'domain'. Cyber is an instrument that, like many others, can be used to support national policies, including ones of deterrence and, more broadly, coercion. One nation may try to deter others from using cyber instruments or threaten to use cyber to deter others from undertaking a range of activities, including but not limited to cyber. The heavy reliance of U.S. military, government, and society on cyber presents new vulnerabilities, opportunities, and complications for international security policy, but the changes in technology do not change first principles of conflict and deterrence. Countries or alliances (sometimes in league with hackers) struggle and cooperate with each other with multiple instruments at their disposal.

Second and perhaps more controversially, all recent and current conflicts with major adversaries are influenced by cyber, even if no attacks or threats are made to networks or by cyber weapons. Not only is computer network exploitation (CNE) widespread and presumably a significant influence on the behavior of states (and at least some non-state actors), but decision-makers know that an enlarged conflict could lead to cyberattacks, and this is very likely to affect their behavior, although exactly how depends on a many factors. The fact that the world exists in a cyber era is an inescapable aspect of security politics.

Third and related, although there are a number of projects devoted to 'cross-domain deterrence', in the broadest sense, there is nothing new here. Contrary to Robert McNamara's famous statement, thanks to the danger of nuclear escalation even more than to explicit threats, nuclear weapons did more than deter nuclear use on the other side. During the Cold War, the realms of

nuclear, conventional, sub-conventional, and space (after satellites were developed) were not segregated from one another, but rather mixed in the cauldron of international politics. (Today, it might be noted, space and cyber are so deeply intertwined that concentration on the latter may lead to important vulnerabilities and opportunities, as well as to decision-makers' floundering in a crisis more than they need to because of the lack of understanding of what is happening.)

Just as deterrence existed before Hiroshima, to take the title of George Quester's fine book on this subject (1966), and yet nuclear weapons both strengthened and changed the links between politics and military instruments, so too the advent of the cyber era both leaves some fundamental principles of deterrence intact and introduces new elements. The following section explores the former.

A Reminder of Some Under-Appreciated Aspects of Deterrence

As Thomas Schelling stressed over fifty years ago, to be effective, threats have to be coupled with promises (1960). That is, the U.S. needs not only to convince the adversaries that it will take certain undesired actions if they harm it, but the U.S. also must convince those adversaries that it will not do so if they comply. If adversaries believe that they will be punished no matter what they do, the U.S. has lost its influence. As noted below, cyber instruments raise particular challenges in this regard; but the general point is that while promises not to do harm are usually pretty credible in day-to-day politics, the task becomes more difficult in a crisis when each side seeks military advantage and knows that the other is similarly motivated.

The conflicting incentives during a crisis are most sharply brought out by the question of whether one government, such as the United States, should attack an adversary's command and control systems. Although this dilemma is well known, it is likely to be heightened by cyber instruments because it is likely to be harder to credibly promise to remain restrained.

Over the longer run, the need for promises in Russian-American relations is linked to Putin's fear of a 'color revolution', which he believes would be sponsored by the U.S. This Russian vulnerability both gives the U.S. potential leverage and creates a danger. The leverage is that Putin might make concessions in order to induce the U.S. to refrain from seeking his overthrow (although this would probably require the U.S. to refrain from actions it considers legitimate and appropriate, such as applauding Russian civil liberties proponents), but the problems are that American promises are not likely to be believed and, more dangerously, that any internal unrest is likely to be attributed to American actions. Putin's fears could then lead him to over-read American pressures, sanctions, and limited military moves in a crisis.

Credibility of threats is, of course, central to successful deterrence, and rivers of ink have been spilled on the subject. Here, the author focuses on only two points that, while known, often get overlooked. First, threats need not be completely credible in order to be effective. In the bad old days in New York when people were robbed on an annoying number of occasions, those individuals did not have to be completely sure that their interlocutor would use force if they refused to comply with his or her demands. Even a fairly low probability would do the trick. Of course, much depends on the target's intensity of motivation to resist and whether he/she believes that giving in to one demand will lead to even greater ones. If an individual is carrying his or her life savings, a much higher estimate of the probability that an adversary is not bluffing

will be required than if what the person were turning over to him or her is of less value. In any case, probabilities figure significantly into decision-making and reactions.

Second, those probabilities have to be calculated by the target. Credibility is not an objective, nor is it a property of the person or state making the threat. Rather it is 'owned' by the target. Some threats can be missed entirely. For example, in the fall of 1969, President Nixon ordered a nuclear alert in order to impress the USSR with the dangers of allowing the war in Vietnam to continue. Unfortunately for him, until the end, the Soviets did not notice. If the target does notice, it may misinterpret. When the Soviets finally realized that something odd was happening, they were puzzled and thought that the Americans were perhaps reacting to the danger of a war between China and Russia. If adversaries have greater understanding, they still may rate the threat as more credible or less credible than the acting state believed it would or that an 'objective' observer would rate it.

Another basic point in deterrence theory is that both sides are in a situation of strategic interaction, in which each tries to anticipate what the other will do, both knowing that the other is engaged in this anticipation. One implication is that each side has to try to estimate what the adversary will do in response to its actions. In a number of past cases, American planners and top officials have failed to do this. Of course, there are real limits as to what is possible because the world is not entirely predictable. Nevertheless, any sensible plan should include a section on what responses are most likely—and, related, what can be learned about the adversary from how it does respond.

Expectations are also crucial in a different way, one that is even more often neglected. When a U.S. adversary takes a significant hostile step, if it has planned adequately, it will expect some response from the U.S. and others. If the U.S. is to alter the adversary's course of action, it will then have to do more than what was expected, or perhaps something different from what was expected. To take only those measures that the adversary anticipated cannot compel it to retract or even deter it from going further.

Of course, this thinking places great demands on intelligence to estimate what the adversary is expecting. In many cases, these demands are simply unreasonable, and analysts will have to make do with informed judgements that are based on less-than-solid evidence. But it is important to make these judgements explicit and to factor them into the process for making recommendations. Without a sense of what the adversary expects the U.S. to do, the adversary cannot make a sensible estimate of the likely impact of U.S. policy.

Intelligence is central to deterrence in a broader sense as well. The U.S. needs to know—or at least be able to estimate—how the adversary sees it, how that adversary will likely interpret U.S. actions, and what that adversary values. At times, what the U.S. has believed to be a potent threat actually was an inducement. For example, during a period of unrest in Cuba in the early twentieth century, the U.S. sought to deter the rebels from attacking American property by telling them that if they did this, the U.S. would respond with force. Unfortunately, however, the rebels believed that American intervention was in their interest, so they gleefully burned American-owned fields and factories.

This might be termed the deterrence trap, and it could be at work in the current era. The leaders of both Russia and China have to worry about maintaining popular support, and an obvious way to do so is to rally nationalism and blame the U.S. for any domestic problems. It is possible, then, that Western economic sanctions against Russia in the Ukrainian crisis may have played into Putin's hands. Hypothetically, if the Russian economy were to falter to the point that there were power outages and food shortages in a subsequent confrontation, American actions (either in the form of economic sanctions or cyber interventions) might be welcomed because the regime could then attribute all the problems to the machinations of the West.

Indeed, in retrospect, the Japanese might have considered the U.S.'s stationing the fleet at Pearl Harbor to be laying a trap. Their attack not only unified the U.S. in fighting World War II, something that otherwise would have been difficult, but also sank many U.S. battleships in shallow waters (from which they could be eventually retrieved), and attacking them spared the aviation fuel-storage facilities whose destruction would have done more to cripple the American effort over the next years.

Just as U.S. deterrence strategy requires a correct reading of the adversary, so understanding the adversary requires the U.S. to understand how the adversary sees it, a task that is at least as difficult. The U.S. is likely to believe that the adversary sees it accurately, or rather that he or she sees it as the U.S. sees itself. This, in fact, is rarely the case. In some previous cases, such as the move to the Yalu in 1950, American leaders failed to understand the extent to which actions that they regarded as unthreatening were seen as threats to the adversary's vital interests. The U.S. is often seen as more menacing than it believes. In other cases, deterrence will be undercut by a U.S. inability to understand how the adversary thinks the U.S. will respond to its aggressive moves. For example, one reason American deterrence failed in 1941 was that the Japanese leaders believed that, in response to initial setbacks at the Japanese hands, the U.S. would be willing to settle for losing a limited war rather than engaging in a protracted and bloody all-out struggle. The Japanese misunderstood the U.S., something the U.S., in turn, failed to understand.

It is not only the adversary that can fail to perceive what is of greatest value to the U.S. In the press of a crisis with its twin imperatives to avoid escalation and secure a favorable political settlement, the U.S.'s own focus may narrow to tactics and the specific situation. Yet, in any conceivable conflict with Russia, NATO is likely to be the greatest prize, with Russia seeking to weaken if not destroy it and the U.S. having the opposite interest. Unlike the simplest situation so dear to the hearts of academics, the confrontation is not purely bilateral, and the NATO countries if not others throughout the world are central to the U.S.—and more important than the issue or territory that is the immediate stake in a crisis. The magnitude of the difficulties confronting the military in making recommendations for effective actions and responses during a crisis, however, is likely to be sufficient to divert attention from this goal.

As a crisis appears and develops, there are also likely to be tensions between the advantages and dangers of moving quickly. Declaring red lines and moving military forces needed to enforce them have obvious advantages from the standpoint of deterrence. During the Cold War, much American policy was built around the perceived efficacy of commitment—staking American reputation on defending positions even when they were militarily indefensible, as in West Berlin. Two points must be made on the other side, however. First, a great deal of the information

received at the start of a crisis is wrong. Although waiting for more information does not guarantee a good understanding of the situation, quick moves are more likely to be based on false understandings. Second, political leaders usually want to keep open as many options as possible for as long as they can and to preserve their decision space. Military planners and leaders are more likely to be impressed by the advantages to be gained by moving quickly and decisively. The possible use of cyber instruments, both by the U.S. and its adversaries, is likely to increase the pressures for speed. A crisis involving the actual or potential use of cyber instruments is also likely to generate unusual if not unique challenges to the government because these would be unprecedented, and top leaders and advisers in the White House (and perhaps in the Pentagon as well) would be less than fully familiar with the cyber world.

Relations with allies, noted earlier, have another aspect: too great a commitment to them may allow them to take advantage of the U.S., either by not contributing to the common defense ('free-riding') or, more dangerously, engaging in actions that they feel are in their interests, but are not in the interests of the U.S. ('moral hazard'). The best known example is that if Taiwan were confident of unconditional American support, it might declare its independence. Similarly, if the Baltic states were certain that the U.S. would and could protect them, they might be tougher on the Russian minorities and trigger a crisis.

Aspects of Deterrence That Are Particular to Cyber

Everyone would agree that the use or potential use of cyber instruments introduces some new elements into deterrence, but, after that, disagreement and uncertainty abound.

The greater role for uncertainty is itself important, however. This is not to imply that confrontations in the Cold War unfolded in fully anticipated ways—if they did, they never would have started. But some degree of shared understandings and routinizations developed, partly as a result of a fairly good idea of what weapons, both conventional and nuclear, could do, and partly because the interactions were repeated. Perhaps the most obvious current uncertainty is exactly what the deployment of any cyber weapon would do. Granted, the U.S. has learned quite a bit through CNE and the non-trivial number of computer network attacks (CNAs); but partly because each network is different, enormous uncertainties apparently remain about the effect of possible cyber moves. For example, it is generally believed that when the Israelis used cyber to blind or disable Syrian air defenses around the North Korean nuclear reactor to pave the way for their air attack, they also turned out some of the lights in Damascus. The problems here are two-fold. First, the U.S. could not know the physical, let alone the psychological and political impact, of exercising various cyber options; and, second, it is likely that the country that is the object of the attack would assume that any effect was the intended one. The obvious danger is one of unintended escalation, but the happier possibility of escalation leading the other side to make greater concessions should not be neglected. For example, one reason Khrushchev ended the Cuban missile crisis was that he had not anticipated that a Soviet air defense commander might shoot down an American reconnaissance plane without authorization from Moscow, and Ayatollah Khomeini made peace with Iraq in part because he believed that the American shooting down of the civilian airliner in 1988 was an indication that the U.S. was about to enter the war. In reality, this shooting was an accident.

But the uncertainties with both the physical effects and the interpretations of cyber instruments are likely to be even greater than in previous eras. Furthermore, in most organizations, uncertainties tend to get played down if not filtered out as things move up the hierarchy. In a crisis, the people at the top will feel unwarranted certainty about what the physical if not the political effects of cyber instruments would be. Another layer of complexity could be added by the discovery of on-going cyber activities. In a confrontation between adversaries, both sides will examine their own networks with great care, and might discover adversary penetrations. Even if these were of a long-standing nature, the sudden realization of their presence would be deeply unsettling.

Needless to say, the inability to distinguish CNE from CNA is a great source of uncertainty. To take just one manifestation, if either side has penetrated the other's command and control system for the purpose of gathering intelligence, it will be very hard for it to reassure the other that it will not trying to convert this access to a disabling attack. Although this article does not explore this point in depth, it is not because it is unworthy of further comment. Instead, the author defers that discussion to others, whose deep involvement with cyber makes them much more knowledgeable than he about this topic.

A related source of ambiguity and confusion is that there is no consensus on the meaning of escalation with cyber instruments. Although the extent of shared understandings about Herman Kahn's 'escalation ladder' during the Cold War (1968) may be exaggerated, everyone would agree that the use of nuclear weapons was on a much higher rung than conventional force. But where would cyber instruments fit compared to others, and what sorts of cyber operations would be more severe, and which would be milder? It is unlikely that anything approaching complete agreement exists about a cyber escalation ladder within the U.S., let alone between the U.S. and possible adversaries. So far, there does seem to be a general assumption of symmetry in that states have replied to cyber attacks with a cyber response (when they have responded at all). This pattern appears to have been the case in the exchanges between the U.S. and its allies and Iran. The fact that the U.S. has multiple weapons in its arsenal, however, means that it can utilize other forms of response. Its use of economic sanctions in response to North Korea's attack on Sony presumably followed from the lack of appropriate targets on the other side, and the use of judicial indictments against Chinese officials, a move of largely symbolic significance that probably stemmed from the desire not to escalate the conflict. What may be most important is what is taken for granted: so far no state has made a kinetic response to a cyberattack. Even though a small bombing raid might in some sense be a de-escalatory reply to a large cyberattack, leaders instinctively prefer to keep the two realms separate, at least for the time being. The old distinction between vertical and horizontal escalation may be blurred or lack utility here. Physical location has less meaning in cyber, although the basic idea of responding to an adversary attack in an area in which it is weak (and the U.S. is not) still makes sense.

Much of this article implies only two sides, but, in fact, there would be multiple audiences. Not only is no state completely unified, but the perceptions of numerous third parties are also important. What would seem like an under-reaction to some allies, for example, could be seen as a dangerous over-reaction by others.

The attribution of CNE or CNA is usually at least somewhat ambiguous, unlike most uses of force (but not all—as the Little Green Men are a reminder of). This both gives the attacker added options and enlarges the role of ambiguity. Presumably, a cyber intervention would be seen as sending a stronger signal if the attacking state takes credit for it. But this does not mean that attacks for which the state does not take credit will be seen as coming from hacktivists or third states. Instead, it is likely that, especially during a confrontation, Russian leaders would attribute any untoward cyber activities to the U.S., even if an accident is really what brought down part of their power grid.

Many military instruments present a tension between secrecy and deterrence, but this is heightened in the cyber realm. A state that has developed an effective weapon or strategy might be tempted to reveal it in order to bolster deterrence but will pay a high price for doing so if the other side can emulate or counter it. This is true for the use as well as the development of forces. The classic example is the Chinese entry into the Korean War. As the U.S. took the offensive, the PRC at first tried to deter it from crossing the 38th parallel by threatening to intervene. But at a certain point, the value of military advantage that was provided by surprise tipped the balance in favor of keeping its deployment into North Korea secret. This trade-off is heightened in the cyber realm. For a state to reveal that it has penetrated or can penetrate various networks or do various kinds of harm can bolster coercive threats, but such displays greatly weaken the weapon. On the other hand, the efficacy of many cyber instruments would be enhanced by leading the adversary to underestimate its vulnerabilities, quite the opposite of what is called for by deterrence. To say anything more would outrun the author's knowledge, and there may be ways to mitigate this tradeoff. The main point here is merely that this dilemma is more central to cyber than to other instruments, and that it calls for careful consideration before a confrontation occurs.

A related tension exists between the desire to prepare for escalation if it occurs on the one hand and the need to minimize the risk of increasing the danger of things spiraling out of control on the other. This potential is especially severe when there are perceived military advantages in the quick use of some instruments. Although experts have convinced this author that glib statements about cyber conflict moving at the speed of light are simply incorrect because the effective use of most instruments depends on careful and lengthy preparation, it is also possible that, in many confrontations, moves that would 'prepare the battlefield' would simultaneously yield real advantage (if the conflict were to continue), enhance deterrence, and/or lead to undesired escalation. It is also unclear at what level decisions about deploying or activating these instruments would be made.

Cyber instruments pose two special complications in this regard. First, the adversary's command systems pose particularly lucrative and tempting targets. To disable or even degrade them gives great advantage, but also increases the likelihood of uncontrolled escalation. Indeed, either side's fear that these vital networks are vulnerable could lead either to strike first or to delegate authority to lower levels, which would multiply the number of individuals who could order attacks (both cyber and kinetic) and, thereby, increase the chance of heightened violence. Second, even if essential networks are not stricken, the use of any cyber instruments is likely to at least somewhat degrade communications among leaders in the capital, as well as between them and units in the field. Both the U.S. and its adversaries have become accustomed to working with amazingly secure and well-functioning communication systems. A decrease in the

trustworthiness and efficiency of these systems, let alone a complete breakdown, would add levels of overload and stress to a situation of extraordinary complexity. Even without any attacks, difficulties would be created by the knowledge that they could occur or the fear that information being received is actually the product of adversary spoofing.

Finally, it is worth noting that because the world has relatively limited experience with cyber, anything the U.S. (or others) does helps set precedents for the future. This is why many analysts argue that Stuxnet was a mistake that will harm the U.S. in the long run. Whether or not this is correct, how the U.S. acts and responds to challenges (or declines to respond) will shape others' expectations and sense of what the 'rules of the road' are.

Conclusion

Perhaps the most important point is also the most banal one: current understanding of the cyber dimensions of conflict is at a primitive stage, partly because the technology and, even more, the uses of it by states and private actors are evolving. The future is probably unknowable because it will depend in part on how important actors think and behave. As has been discussed, cyber escalation ladders are quite unclear, and it is up to actors to determine what the order of the rungs will be. Analysts could develop elaborate schemes, but it is far from clear that they would match the way the important actors would see it. Behavior as well as thinking will help chart the future. Although the prospects for cyber arms control agreements seem dim, states may yet be greatly restrained in their use of CNA, although almost surely not in CNE. The rules of the game are being written as it is being played, and what each player does influences not only specific responses by others, but many of the contours that will guide future play.

References

Kahn, H 1968. *On escalation*, Praeger, New York, U.S.A.

Quester, G 1966. *Deterrence before Hiroshima*, Wiley, New York, U.S.A.

Schelling, T 1960. *The strategy of conflict*, Harvard U.P., Cambridge, MA, U.S.A.