Chapter Title: Escalation Management

Book Title: Crisis and Escalation in Cyberspace

Book Author(s): Martin C. Libicki

Published by: RAND Corporation

Stable URL: https://www.jstor.org/stable/10.7249/j.ctt24hrx7.11

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at https://about.jstor.org/terms

This content is licensed under a RAND Corporation License. To view a copy of this license, visit https://www.rand.org/pubs/permissions.html.

*RAND Corporation* is collaborating with JSTOR to digitize, preserve and extend access to *Crisis and Escalation in Cyberspace*

المنارة للاستشارات

www.manaraa.com

This content downloaded from 81.218.45.221 on Mon, 08 Nov 2021 13:47:06 UTC
All use subject to https://about.jstor.org/terms

CHAPTER FOUR

# Escalation Management

Once a crisis has blossomed into conflict, crisis management becomes escalation management. The success of escalation management depends on the fact that both sides would prefer less disruption and violence rather than more of it—but not necessarily before they make their point to one another. At the very least, both sides share an interest in keeping control over what breaks out rather than ceding control to fate, the passions of warriors, the intrigues of factions, or third parties.

Admittedly, escalation in cyberspace remains a speculative topic. Few government officials have declared their red lines. The cyber equivalent of Herman Kahn's *On Escalation*[1] is yet unwritten. Not only do we lack a discrete metric for cyberwar, there is no good way to measure the proportionality of various cyberattacks systematically and consistently (e.g., "this act is more heinous or dangerous than that act").

After a quick review of escalation motives, this chapter covers three topics: (1) the risks of escalation associated with cyberattacks in various contexts, (2) third-party escalation, and (3) the difficulties of controlling escalation using tit-for-tat logic. Afterward, we examine escalation narratives that each side may offer and then issues associated with the C2 of cyberwarriors to implement escalation management. The chapter's context is a conflict in which cyberattacks matter in their own right, rather than being simply one more way to prosecute a target already threatened by kinetic means.

[1] Herman Kahn, *On Escalation: Metaphors and Scenarios*, Praeger, 1965.

المنارة للاستشارات

This content downloaded from
81.218.45.221 on Mon, 08 Nov 2021 13:47:06 UTC
All use subject to https://about.jstor.org/terms

## Motives for Escalation

Those who would manage escalation by exercising self-restraint and persuading adversaries to do likewise should start with a sense of what the other hopes to get from unilateral escalation—that is, crossing some hitherto uncrossed red line.

A primary purpose of escalation is to gain military advantage.[2] Yet, a thinking combatant will recognize that, because escalation begets escalation, the military advantage from escalating will have to trump whatever military disadvantage arises when the adversary does likewise.

Calculating net advantage is tricky. The presumption that the adversary will escalate one level in response to a one-level escalation may fail if the adversary calculates that it loses on that round and thereby raises the stakes.[3] After two rounds, the advantages to the escalating side may disappear while the pain does not. In cyberspace, such calculations are particularly complex. Thresholds have yet to be established, or even described in common words. Worse, although each side can recognize the vulnerability of another after having scoped it, recognizing one's own vulnerability, and hence susceptibility to retaliatory cyberescalation, is inherently difficult:[4] If one were already aware of such vulnerabilities, chances are that they would have been already

[2] For a richer treatment, see Morgan et al., 2008, especially the first few chapters.

[3] Albert Wohlstetter and Richard Brody, "Continuing Control as a Requirement for Deterring," in Ashton B. Carter, John D. Steinbruner, and Charles A. Zraket, eds., *Managing Nuclear Operations*, Washington, D.C.: Brookings Institution, 1987, pp. 142–196, posits a hypothetical conflict with the Soviet Union circa 1985 that attacks NATO's southern flank with nuclear weapons to shatter the alliance. NATO concludes that it lacks a comparably good nuclear target that would have a similar effect, so it escalates to find its own sweet spot, which, by definition, is a sour spot for the Soviet Union, prompting it to counterescalate, and so on.

[4] One can parameterize certain types of vulnerabilities (e.g., the likelihood that a user has a compromised machine) statistically, but many of the nastiest attacks do enough damage if they succeed but once.

المنارة للاستشارات

This content downloaded from
81.218.45.221 on Mon, 08 Nov 2021 13:47:06 UTC
All use subject to https://about.jstor.org/terms

fixed.[5] It is thus easy for one side to argue that the *net* effect of escalation is positive because of inherent asymmetries in knowledge.

A secondary purpose of escalation is to signal seriousness, both to one's own side and to the other. To one's own side, it is a signal of support. A state that sends its military to fight and die in a theater is saying that it is willing to risk the adversary escalating to attacking the homeland in order to pursue military goals in theater. To the other side, escalation can say, "cut it out or someone is going to get hurt"; it can convey, for instance, that cyberespionage has reached a point at which the pain is tantamount to that of a cyberattack. If cyberescalation supports a theater military operation, it may communicate that the outcome of such a conflict matters a great deal.

A third purpose is to demonstrate one's power: "we can do this to your systems despite your best efforts to keep us out; now, do you trust them?" A related purpose is to carry out a contest of pain (or perhaps a contest in risk-taking, per Schelling's argument in *Arms and Influence*[6]). This presumes something called *escalation dominance*—the ability to outmatch one's foe at all levels of escalation.

A fourth purpose is to test the temper of opponents: How far are they willing to go? Are they rational and measured or irrational and erratic? An escalatory move may be tried to see how opponents would react. The advantage of doing so in cyberspace is that it provides *some* insulation against overreaction in the real world. But the value of such a test assumes that the same personality patterns that manifest themselves in cyberwar will manifest themselves similarly in physical (kinetic) war. Those that use cyberattacks to ping the other side have to contend with *four* sources of error if the response comes back by means of a cyberattack: (1) the difference between the intended attack

[5] This assumes that the government can fix vulnerabilities in infrastructure systems it does not own—an unwarranted assumption in peacetime, but plausible in wartime if such vulnerabilities threatened the war effort.

[6] Thomas C. Schelling, *Arms and Influence*, New Haven, Conn.: Yale University Press, 1966.

المنارة للاستشارات

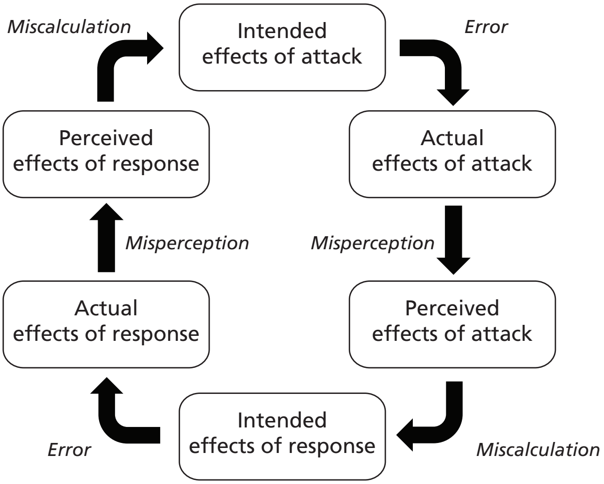
This content downloaded from
81.218.45.221 on Mon, 08 Nov 2021 13:47:06 UTC
All use subject to https://about.jstor.org/terms

and its effect,[7] (2) the difference between the effect and its perception by the target, (3) the difference between the target's intended response and the effect it had, and (4) the difference between the actual effect of the target's response and how it was perceived by the original perpetrator (which is a problem both of measurement and of correlating the response to the original impulse). To this one can add miscalculation on the attacker's part about how the adversary will respond and the latter's miscalculations in response. The real signal may get lost in the noise of all the echoes, as illustrated in Figure 4.1.

## Does Escalation Matter?

Our treatment assumes that states do have a positive interest in controlling their adversaries' use of cyberattacks and are willing to curtail their own use to that end. Here, we pause and ask, how much difference *is* there between a no-holds-barred cyberwar campaign and a

**Figure 4.1**
**Sources of Imprecision in Tit for Tat**


Intended effects of attack
Error
Actual effects of attack
Misperception
Perceived effects of attack
Miscalculation
Intended effects of response
Error
Actual effects of response
Misperception
Perceived effects of response
Miscalculation

RAND *MG1215-4.1*

[7] Or, if one is measuring the response from the "actual" effect of the test cyberattack rather than the intended effect, the error may come from the difference between the attacker's perception of the effect and the actual effect.

المنارة للاستشارات

This content downloaded from
81.218.45.221 on Mon, 08 Nov 2021 13:47:06 UTC
All use subject to https://about.jstor.org/terms

modulated cyberwar campaign? Violent war features very wide bands. Without war, a state's greatest worry about losing its citizens to violence is crime. By contrast, a nuclear-armed peer could kill everyone and break everything. There is a lot of scope for escalation within that band.

Now consider cyberwar. In today's environment, cybercrime is constant, with an annual cost to the United States in the billions of dollars and a plausible premise that, if a system with requisite vulnerabilities has something worth stealing, theft *will* take place and sooner rather than later. Because the general noise level is high in cyberspace, any retaliation that merits notice as such has to be loud. So the bottom is quite high.

The top end may be low, relative to conventional, much less nuclear, war. As noted, no one has yet been killed in a cyberattack, and there is scant indication that a full-blown attack could kill as many as a normal year's flu epidemic.[8] The most–commonly cited worst-case scenarios concern attacks on power companies that succeed in damaging a great deal of equipment, but extrapolating from Idaho National Laboratory's Aurora experiment or even Stuxnet to such a scenario is quite a stretch (safety and control considerations suggest that confused power equipment default to shutting down rather than damaging itself or other equipment).[9] Similarly, there is little basis for knowing how much damage can arise when modern process control and financial systems fail, or how well timely and intelligent human intervention can mitigate such costs. Similarly, there is little evidence of how much operators can damage their own equipment if they are misled by monitors that have been deliberately corrupted. Whether infrastructures have weaknesses that no one has seen yet to exploit but whose effects could be sharp and hard to fix remains unknown. All one knows is

[8] Roughly 6,000 per year, based on Centers for Disease Control and Prevention statistics from winter 1976–1977 through winter 2006–2007 (Centers for Disease Control and Prevention, "Estimates of Deaths Associated with Seasonal Influenza: United States, 1976–2007," *Morbidity and Mortality Weekly Report*, Vol. 59, No. 33, August 27, 2010, pp. 1057–1062).

[9] See Jeanne Meserve, "Sources: Staged Cyber Attack Reveals Vulnerability in Power Grid," CNN, September 26, 2007.

المنارة للاستشارات

This content downloaded from
81.218.45.221 on Mon, 08 Nov 2021 13:47:06 UTC
All use subject to https://about.jstor.org/terms

what has happened so far. By the standards of conventional warfare, the damage has not been terribly impressive—so far.

Perhaps the real reason to control cyberescalation is that matters may not end in cyberspace. One side may see that cyberattacks on targets that *were* off-limits to kinetic attack legitimize a kinetic attack on comparable targets: If cyberattacks on a sensitive system put lives at risk, why are they different from a kinetic attack that puts the same lives at risk? So, should targets considered off-limits from a physical attack also be off-limits from a cyberattack *that offers the potential of similar collateral damage*? A state hit by a devastating cyberattack may conclude that, like Indiana Jones, it is tired of getting cut with cyberknives and whip out its kinetic pistol. Such a reaction would trade the limited risks of cyberescalation with the nearly unlimited risks of violent escalation, but states may take that risk. Although violent escalation is beyond the scope of this chapter, it does present a serious risk that those that would escalate in cyberspace cannot ignore.

## Escalation Risks

Just as deterrence works only if the adversary believes it does, so too with escalation: The adversary's perception of red lines determines whether one's own cyberattacks are escalatory. The fact that adversaries determine what is escalatory sets the context for this chapter's question: What is the escalatory potential inherent in cyberattacks? We examine three contexts: precrisis preparations, operational cyberwar within local conflicts, and escalation beyond operational cyberwar.

### Escalation Risks in Phase 0

A state that faces the prospect of kinetic conflict should anticipate being hit by cyberattacks. Yet, if it views the prospect of conflict as possible rather than inevitable, it faces a choice. Modulating its activities may avoid exciting the other side and contribute to a peaceful resolution, but it may also signal distaste for battle and leave vulnerabilities untended—both of which may encourage a determinedly hostile state.

المنارة للاستشارات

This content downloaded from
81.218.45.221 on Mon, 08 Nov 2021 13:47:06 UTC
All use subject to https://about.jstor.org/terms

Five activities may characterize phase 0 activities in cyberspace: (1) increasing defensive preparations, (2) demonstrating offensive capabilities, (3) accelerating cyberespionage (e.g., to find more vulnerabilities), (4) inserting implants and back doors (e.g., to facilitate attack), and (5) disrupting problematic communication outlets. Each one carries its own type of escalation risk.

Increasing cyberdefensive preparations should carry little escalation risk. Most such preparations are invisible to the adversary. They are generally not adversary-specific. Furthermore, they are also entirely legitimate. Nevertheless, benign outcomes are not guaranteed. Some preparations will be hard to hide (e.g., cutting users off from the Internet). Also, if the state bulwarking its defenses wishes to communicate as much in order to dampen an adversary's temptation to carry out a first strike, some preparations must perforce become visible. The adversary can concede their legitimate nature and think nothing more of the attempt. But adversaries are constantly assessing the intentions of their rivals and may conclude that defensive preparations are being made in order that offensive cyberattacks later be carried out with impunity (as discussed in Chapter Six). A great deal depends on what the adversary takes to be an indication and what it takes to be a warning.

Demonstrating offensive capabilities, by contrast, is both visible and invisible if successful. It may be regarded as hostile, particularly when carried out during a crisis. It is definitely a warning to adversaries not to start trouble, but it may also convince adversaries that trouble is coming. A lot may depend on whether the state that is witnessing such demonstrations believes that it is being coerced to yield or settle. If asked to yield, it may bridle. If asked to settle, it may reason that demonstrating cyberattack capabilities may make a point at the cost of reducing its later effectiveness by hinting at the target's vulnerabilities. Such a conclusion would suggest that a shot across a bow is more of a reminder than a preattack maneuver. Again, a great deal depends on how the adversary thinks.

Accelerating cyberespionage should also be invisible, hence unproductive of trouble. Discovering the odd penetration is no proof that activity has accelerated because cyberespionage is always taking place, unless the discovered penetration affects systems previously

المنارة للاستشارات

This content downloaded from
81.218.45.221 on Mon, 08 Nov 2021 13:47:06 UTC
All use subject to https://about.jstor.org/terms

thought off-limits to attack. Again, a lot depends on the things to which the adversary reacts.

Inserting implants follows the same logic: They should be invisible. Even if implants are found, they may arguably be prefatory to cyberespionage rather than cyberattack. Nevertheless, the adversary may find such implants to be akin to discovering mines in their waters: clearly hostile and putatively an act of war. Although earlier chapters cautioned that finding them ought not necessarily create a crisis, adversaries may not heed such cautions.[10]

Finally, using cyberattacks to disable contentious communication channels, such as web sites that incite to violence, may be a step in favor of crisis resolution or may be viewed as a violation of sovereignty.[11] The helpful conclusion requires that adversaries (1) feel that such web sites were themselves unhelpful but lacked the skills or the political cover to disable such sites themselves and (2) are more likely to let their private relief rather than the public posturing guide their actions.

### Escalation Risks for Contained Local Conflicts

In theory, operational cyberwar—carrying out cyberattacks on targets that are considered legitimate war targets—should not be considered escalatory. It is just another way to accomplish the same end, and with fewer lives at risk. But sometimes, an act is judged escalatory based not on what it does but how it does it (e.g., taking out a bunker with chemical weapons is considered more heinous than doing the same job with

[10] The discovery of one piece of malware tends to increase the odds of finding others of both the same type and different types. The first malware suggests the possibility of a systematic campaign, which, if nothing else, favors intensifying search efforts; greater search tends to lead to more finds. In some cases, the signature of the discovered malware (or the communications to the C2 server) may help identify subsequent copies and even derivatives (much as the discovery of Stuxnet may have somewhat hastened the discovery of Duqu and Flame). Unfortunately for crisis management, it may be hard for the public to distinguish a cascade of discoveries bunched closely in time from a cascade of attacks that are similarly bunched. The actual attacks may have, in fact, been emplaced over a far longer period in the past.

[11] If the affected web site is in a third country, legalities and the reaction of the latter may have to be taken into account.

المنارة للاستشارات

This content downloaded from
81.218.45.221 on Mon, 08 Nov 2021 13:47:06 UTC
All use subject to https://about.jstor.org/terms

high explosives).[12] The Japanese considered the first use of firebombing (March 1945) to be escalatory even though the attack on Dresden, Germany, had already taken place. The use of cruise missiles in Bosnia (1995) was considered escalatory.[13] Although it is unclear whether such sentiments were anything more than sentiments (because neither target could escalate in response), the broader point stands.

If opponents believe that cyberweapons have mysterious effects, their use will be seen as escalatory even though, measured in terms of actual effects, they should not be. Adversaries may also convince themselves that, although the cyberattacks per se were in bounds, their use against military targets portends their use against civilian targets because the latter can be surreptitiously attacked via cyberspace even if kinetic attacks on them would be universally considered off-limits. Again, it depends on what adversaries think.

### Escalation Risks for Uncontained Conflicts

Cyberescalation beyond the immediate local conflict can go down one of several paths,[14] and each path carries its own escalation risk.

One path is to attack systems with effects beyond the conflict zone. Thus, attacks on a system that supports local combat operations may disable the adversary's ability to carry out other operations. Such systems could physically sit in the theater or, alternatively, out of the theater; in cyberspace, physical location is almost an afterthought. Although legitimate targets of war and cyberattacks on them ought not be considered escalatory, the adversary's perspectives are what matter.

Another path is to attack systems that have civilian uses. Some of these may be systems that control homeland assets that are used to support a war (e.g., a cyberattack on the management of a military port in the homeland). Further along the path is an attack on dual-use facili-

[12] Using means rather than ends as the measuring rod of escalation downplays the possibility that means may be shifted, not to gain an advantage but because prior means have been rendered unavailable. Cyberattacks, for instance, may be used as an attempt to replace effects that electronic warfare previously offered.

[13] Morgan et al., 2008.

[14] A similar point is made in Kahn, 1965, p. 5.

المنارة للاستشارات

This content downloaded from
81.218.45.221 on Mon, 08 Nov 2021 13:47:06 UTC
All use subject to https://about.jstor.org/terms

ties (e.g., that port supported both commercial shipping and military logistics). Even further along the path is an attack on a primarily civilian activity: a power plant that supplies a city but also an integrated air defense system. A larger step by way of escalation is to attack a facility *with the intent* of persuading civilians to pressure their government into suing for peace. Such attacks are more likely than the purely military attacks to be perceived as escalation.[15] How the other side creates narratives around such attacks may determine what kind of response may be forthcoming. One approach is righteous wrath: The cyberattacker has escalated a local conflict into a global conflict, and all restraints are off. Another would use the attack on the homeland to mobilize its population to support the remote conflict but confine the response to the theater. Or the adversary, unwilling to escalate a local into a global conflict, can just shrug the attack off.

There are escalation steps even within the category of coercive attacks. Attacks on infrastructure are bad, but attacks that disable or disrupt safety systems (e.g., air traffic control) are worse, and those whose sole purpose is to create civilian casualties (e.g., hospital medication monitors) are worst. The closer cyberattacks get to civilians, the more likely they are to violate the laws of armed conflict and UN treaties. Finally, strategic attacks imply that states *can* be coerced, which is insulting and not just injurious (whereas it is no insult if other states try to disarm a state).

The third path entails attacks on systems that portend wider conflict (e.g., on strategic warning systems or, more broadly, corruption attacks that may make the target worry about the quality of its over-

[15] Would an attack on an information system, such as a cloud host, situated in the territory of an innocent third party but critical to the conduct of the adversary's campaign be considered out of bounds? Perhaps that question is premature. First, militaries are more reluctant than commercial enterprises to entrust their critical information to neutral third parties even if their information is encrypted. Second, only some of the ways of attacking such information systems are attacks on someone else's "territory": Client-side corruption probably would not count, and it is unclear whether exploiting a flaw in the server's software to corrupt data content would count either; conversely, disabling a particular server might raise third-country issues especially if the server has other customers. Might matters be modulated if the third country were told of the risks assumed by its hosting databases that support a third country's wars?

المنارة للاستشارات

This content downloaded from
81.218.45.221 on Mon, 08 Nov 2021 13:47:06 UTC
All use subject to https://about.jstor.org/terms

all C2 over both fighters and weapons). Similarly, crippling systems that hamper the target state's ability to maintain its hold on power may be misinterpreted as prefatory to a regime-change campaign. The same holds for disruptive but especially corrupting attacks on state-friendly media and internal security systems. If the adversary is nervous enough about internal stability, then a cyberattack on the capability of its domestic security forces may trigger a panicked and potential escalatory response (unless such attacks persuade the adversary to back off and conclude that bigger stakes than information security are on the table). Similarly, something like the (so-called) Great Firewall of China would be off-limits, despite how richly apropos a target it may appear. A related set of attacks to avoid is one that undermines the basic trust that citizens have in their government and comparable institutions (e.g., corruption attacks on the financial system). Systems behind which the adversary has put public prestige—perhaps because they allow it to fulfill an important promise or because they have been touted as secure—may also force the attacker to escalate.

The effect that strategic cyberwarfare can have on the narrative of conflict also has to be considered. A state whose conflict goals are local and definite may unwittingly create another narrative by escalating into the other side's homeland. It may aver that the purpose of such attacks was operational in that the target systems directly supported the adversary's war operations, or coercive in that it expected the population to demand that the local conflict be brought to an end. The besieged state may justifiably conclude that the purpose of these attacks was regime change and react as if the stakes had changed. It might work: Attacks on Belgrade and on facilities owned by friends of Milošević may have convinced the regime that losing Kosovo was better than losing everything,[16] but the risks of a response should not be overlooked.

[16] Stephen T. Hosmer, *The Conflict Over Kosovo: Why Milosevic Decided to Settle When He Did*, Santa Monica, Calif.: RAND Corporation, MR-1351-AF, 2001.

المنارة للاستشارات

This content downloaded from
81.218.45.221 on Mon, 08 Nov 2021 13:47:06 UTC
All use subject to https://about.jstor.org/terms

## Managing Proxy Cyberattacks

Proxy cyberattacks may well be a feature of future wars if and when many states acquire the requisite offensive cybercapability and their targets acquire systems that are simultaneously important to warfighting and vulnerable to attack. Third parties may have all manner of reasons to jump in. They may wish to weaken one side or another's ability to carry out military operations. Perhaps, they would like to see the conflicts of others grow harder to withdraw from thereby letting such third parties wreak mischief in other neighborhoods with greater confidence they will not be interfered with. Such attacks may also be carried out as a live-fire training exercise or as active pinging—a way to collect intelligence that passive methods cannot offer. The attractions of third-party meddling are enhanced by the reduced likelihood of getting caught: Not only are multiple parties wreaking mischief at the same time, but each combatant's tendency would be to blame cyberattacks on its battlefield foes rather than on third parties.

To the extent that proxy cyberattacks matter, each party to a conflict may have to think about how to suppress such attacks without creating new escalation challenges. In this section, we examine two scenarios: (1) when third parties are covert and (2) when their participation is overt.

### What Hidden Combatants Imply for Horizontal Escalation

A two-party conflict may easily become a multiparty free-for-all in cyberspace, making attribution more difficult and creating difficult decisions about how to recognize and respond to third-party attackers. To illustrate as much, consider a cybercrisis between the United States and Iran that arises from a politico-military crisis.

So far, the contest seems simple: the United States versus Iran, with the prospect that each will carry out cyberattacks on the other. More to the point, the first suspect in any attack on U.S. forces will be Iran and vice versa. Assume, for the sake of discussion, that the United States has self-imposed limits on its own cyberattacks (e.g., it will not attack civilian targets unless necessary to hinder Iranian military capabilities or operations).

المنارة للاستشارات

This content downloaded from
81.218.45.221 on Mon, 08 Nov 2021 13:47:06 UTC
All use subject to https://about.jstor.org/terms

Iran, in this scenario, however, may well have multiple targets for its cyberwarriors. They include the U.S. military and anything that can annoy the United States (unless they think that an enraged United States is a more dangerous foe). But would Iran stop there? For historical reasons, the Iranians tend to blame the United Kingdom more than a neutral reading of Iran's circumstances would warrant: Such attacks may be meant as punishment for real or imagined offenses since, but may also be meant to discourage possible UK involvement. Other potential targets include Sunni Arab states that have made no secret of their fear of Iran (and that may be inclined to help U.S. kinetic and cyberforces). If Iran follows Saddam Hussein's logic from the first Gulf War, it may eye Israel as a target as well as a way of goading Israel into doing something that may alienate its Sunni Arab foes.

Conversely, it is by no means obvious that those Iran would target are waiting patiently to be attacked before they respond. Iran's foes may figure that a cyberattack on Iran would help U.S. efforts. If the United States has, in their view, unwisely retrained its own operations, it may hope to goad Iran into striking nonmilitary targets of the United States by striking corresponding targets within Iran, thereby deepening the U.S. commitment.

Such third parties would be a minor problem compared with what would arise should a seriously competent cyberpower (e.g., Russia or China) get into the fray. Cyberspace permits such powers to curry favor with one side without necessarily making the other side an enemy—something that would be very difficult for combatants in the physical world, where attribution is more, albeit not perfectly, obvious (in that sense, carrying out cyberattacks has many of the same attributes as lending support by providing intelligence). Such third parties may also have a stake in starting or, conversely, halting a crisis: If the crisis turns into conflict, they have a stake in the outcome. A last motivation for outside powers is to find out where the U.S. military is vulnerable to a disruptive attack, as well as how the U.S. military would respond to an attack. A grateful Iran would be more than willing to supply them intelligence on U.S. forces of the sort that could be gained only by being in hostile contact with them. Iran can also lend them platforms from which to test attacks that require being within range

المنارة للاستشارات

This content downloaded from
81.218.45.221 on Mon, 08 Nov 2021 13:47:06 UTC
All use subject to https://about.jstor.org/terms

to U.S. radio-frequency (RF) networks. If there are two such powers, Iran could play one off against the other. Normally, this interest could not be pursued, given the consequences of getting caught, but, if one combatant will be predisposed to blame its adversary rather than a third party for any mischief in cyberspace, it may figure that the risks are lower (conversely, the victim may make a point of warning third parties away from interference by threatening harsher reprisals and an itchier trigger finger precisely because third parties create such troubling issues).

With these dramatis personae at play, how can the United States navigate in these treacherous waters without unnecessarily broadening its conflict? A cavalcade of cyberattacks, failures, misread results, collateral damage, cascading effects, narratives of power, accusations, overconfident attempts at attribution, retaliation based on such attribution, and counterretaliation are all possibilities.

What would the United States do with knowledge that Iran is getting help? Perhaps it would be in the U.S. interest to "discover" that Iran had carried out the more-sophisticated attacks if it solidifies domestic support for military operations. It may also be easier to convince everyone to take cyberdefense more seriously if they believe that a middling power, such as Iran, could carry out sophisticated cyberattacks.

True, such an approach would hardly discourage major cyberpowers. Yet, how badly should the United States want to discourage them? Having them attack U.S. forces throws a spotlight on what they can do; there is intelligence to be mined there. Unfortunately, as noted, it also gives *them* a fairly good hint about what U.S. forces can do—and so there is intelligence for *them* to mine there. Who learns more quickly? Can the United States usefully deceive others about its capabilities better than they can deceive the United States about theirs?

Otherwise, how could third parties be persuaded to stop? First, they would have to be convinced that the United States knows they are up to no good rather than believe that the United States is casting about for someone (other than Iran) to blame because the going is rougher than expected. Complaints need to be credible. Second, they would have to believe that the United States could put sufficient lever-

المنارة للاستشارات

This content downloaded from
81.218.45.221 on Mon, 08 Nov 2021 13:47:06 UTC
All use subject to https://about.jstor.org/terms

age on them, either through sub-rosa channels or by taking the chance of going public and doing something before the entire world. If the United States does go public, will third parties deny their participation and argue that the United States is just whining? If their denials are absent or at least insincere, will they back down or conclude that, having been so accused, in for a dime, in for a dollar? If the latter, would they support Iran more overtly—say, with intelligence or equipment—thereby complicating U.S. efforts? In today's environment, in which Iran is the most powerful country that does *not* value stable relations with the United States, denial seems the more likely, unless the United States really pushes the matter. Either way, the U.S. gains from acting on its knowledge may be mixed.

Incidentally, this scenario should illustrate why horizontal escalation, the successive entry of the uninvolved into a war on one or both sides (or how World War I started), is of lesser concern with wars in cyberspace. It is difficult to know who is *not* a combatant in cyberspace at any point in time. Furthermore, the entry of others may not matter nearly as much as it does in conventional conflict, in which numbers matter: One state that joins its forces with another to fight as one can tip the battle. In cyberspace, arithmetic superiority does not mean the same. True, two entities combining their search for vulnerabilities in the same target are likely to be more rather than less efficient, but only if they coordinate their efforts correctly.[17] The likelihood that such cyberwar entities work in nonmutual compartments suggests that this is less likely. Furthermore, given the likelihood that the roster of unexploited and accessible vulnerabilities in the adversary get slimmer after the initial cyberattacks, synergy requires that the two partners be working together *well before* conflict has started, which, by definition, is not escalation.

[17] Adding the forces of one to the search agenda for the other, conversely, may not be as efficient as having each partner pursue its own approach separately, *if* a failure in imagination rather than a shortfall of effort better explains why attempts to penetrate an adversary's system falls short.

المنارة للاستشارات

This content downloaded from
81.218.45.221 on Mon, 08 Nov 2021 13:47:06 UTC
All use subject to https://about.jstor.org/terms

## Managing Overt Proxy Conflict

Proxy war may also take place when a state with sophisticated cyberoperators openly supports one side in a local war. Even if outsiders play by Las Vegas rules (what takes place in-theater stays in-theater), information systems span the world. The mischief perpetrated from outside the theater can affect systems in theater and vice versa. In physical combat—using the Korean and Vietnam wars as examples—the bounds between allowable and proscribed targets were mostly observed. Chinese forces were fair game for U.S. forces below but not above the Yalu River. Russians avoided the Korean theater except for (possible) air combat. U.S. forces were not attacked out of theater. During the First Indochina War, the United States was liberal in sending France supplies, but not people. In the Vietnam War, similar rules applied: In theory, Russian and Chinese "advisers" to North Vietnamese forces manning Russian or Chinese equipment, mostly SAMs, were not explicitly off-limits. Yet, some U.S. policymakers worried about unintentionally killing them (while others were disappointed that they escaped harm).

Are Las Vegas rules possible? Will cyberwar assistance be considered akin to supplies or forces? The fact that cyberwar involves people says forces, but the immunity of cyberwarriors sitting out of theater makes it look more like supplies. Local hackers may be trained on and supplied with exploit tools, information on vulnerabilities, and intelligence on targets. After all that, figuratively pulling the trigger may add very little to culpability.

The links between a local combatant's and its great power friend's systems may color whether friends of each side come to blows. Can systems operated by the local combatant be attacked without interfering with systems of its great power friends? Are the systems the friend brings into theater densely connected to its own global systems? If one side's friend harms the in-theater systems of the other side's friend, would the latter want to make an issue of it? Can the attacker's friend argue local military necessity? Can the target's friend retort that the attack was meant to harm it directly and not influence the local fight? Physical boundaries of the sort that help distinguish acceptable from unacceptable behavior are not as reliable a guide in cyberspace, so the

المنارة للاستشارات

This content downloaded from
81.218.45.221 on Mon, 08 Nov 2021 13:47:06 UTC
All use subject to https://about.jstor.org/terms

usual firebreaks do not exist. One can imagine a continually escalating confrontation that, at some point, requires either negotiations of some sort to establish a new and less obvious firebreak or, failing that, calls for one or the other party to back down unilaterally, lest general war in cyberspace ensue.

So what norms should apply? In some cases, physical boundaries may, for lack of a better alternative, stand in for cyberboundaries. Systems that sit outside the war zone are off-limits to a cyberattack even if they help the local combatant fight, just as supplies warrant a similar status. However, the same would not apply to in-theater portals to such a system. Hence the question: How much should an attacker be expected to know about how local systems and access points are connected to global systems of the great power friend?

Potential asymmetries plague the application of any such norms. If, on one side, local combatants and its global friend kept a good wall between their systems, but the other side does not, then attacking the one side's local systems would carry less risk of escalation than attacking the other side's local systems. Why should the latter get a free pass just because of its architecture? Such asymmetries are compounded by ambiguities in cyberspace. If the citizens of one side's friend depend on capabilities that go haywire if those of its local combatant ally are hacked (such systems could easily sit in third countries), and the other side attacks and claims that its attacks were legitimate, will the other side be seen as credible or as opportunistic?

Avoiding escalation in such scenarios might require such great powers to carefully separate their global systems from those sent to theater and require attackers to exercise great caution to ensure that their cyberattacks have precise effects—never easy, even under the best of circumstances. But it would not hurt for either side to realize that accidents happen, especially in war zones.

## The Difficulties of Tit-for-Tat Management

In 1980, after running a set of extended prisoner's dilemma contests, Robert Axelrod concluded that a tit-for-tat strategy was the optimal

المنارة للاستشارات

This content downloaded from
81.218.45.221 on Mon, 08 Nov 2021 101 Jan 1976 12:34:56 UTC
All use subject to https://about.jstor.org/terms

one.[18] Tit-for-tat strategy is simple: Do not start a fight; if hit, hit back on the next turn; if not hit on a turn, do not hit back on the next turn. The strategy's extension to escalation is straightforward. Not for nothing do states respond to escalation with escalation of their own in the justified belief that such a strategy is best suited to ensure that no one escalates. Hence, intrawar deterrence (the threat of counterescalation as a way of inhibiting the escalation of combatants).[19]

Yet, the extension of such a strategy to cyberspace is problematic. The case for tit-for-tat strategy assumes that intent equals effects equals perceptions. But cyberspace is sufficiently noisy that tit-for-tat strategies may have harmful effects. The problems of intrawar deterrence may be as daunting as the problems of deterrence overall.

### The Importance of Preplanning

Cyberattacks, particularly against hard targets, require considerable scoping of the target. Those who wrote the Stuxnet worm, for instance, took many months understanding the relationship between the Siemens programmable logic chip for which the worm was written and the Iranian centrifuge plant whose operations it was trying to hinder. Planning for conventional strikes is more straightforward and typically much quicker, particularly if there are no worries about getting the delivery vehicle home safely.

The need for prewar planning carries implications for escalation management. If not done, the list of targets that can be struck immediately will be correspondingly reduced. Most of the easy targets will be those that are easy because they are not important, hence not well defended. However, some of the easy targets may be those that were

[18] The problem and the strategy, developed by Anatol Rapaport, are discussed in Robert Axelrod, *The Evolution of Cooperation*, New York: Basic Books, 1984. The term *prisoner's dilemma* describes a situation in which each of two players (prisoners) must choose whether to compete with (by ratting on) or cooperate with (by staying silent about) the other. Each player's individual advantage lies with competing with the other (whether or not the other player competes or cooperates), but both would be better off if they both cooperated.

[19] Intrawar deterrence consists of threats against acknowledged adversaries as a way of limiting the depth, breadth, or frequency of their attacks; interwar deterrence is meant against those that have yet to attack.

المنارة للاستشارات

This content downloaded from
81.218.45.221 on Mon, 08 Nov 2021 13:47:06 UTC
All use subject to https://about.jstor.org/terms

not particularly well guarded because their owners did not conceive that anyone would profit from attacking them. Thus, hospitals tend not to be the most security-conscious institutions, compared to say, banks.

If commanders want to escalate and have not prepared the cyberbattlefield, their options are limited, leaving mostly targets whose disruption or corruption would have low and hence unimpressive impacts or those that have high impacts by virtue of their shock value. Unfortunately, shock value is not conducive to escalation management.

Thus, it helps for a state to think through its possible target set in advance.[20] It may decide to put certain targets off-limits and therefore not scope them, but it cannot change its mind instantly.

As a corollary, a cyberattack that fails to elicit a retaliatory response may be interpreted as one that did not cross the other side's red line. The truth may be that the victimized state, surprised to be attacked in that way, had simply not developed a capability to respond in kind.

### Disjunctions Among Effort, Effect, and Perception

A tit-for-tat strategy that works well in a quiet environment may not work so well in a noisy one. An important problem arises from the potential discordance among intentions, effects, perceptions, and announcements. As noted, predicting battle damage is extremely difficult. Facing that problem, those that would escalate may try a shotgun approach, hoping that something will break. By doing so, they effectively renounce any precision in escalation management. They also give up trying to make a point by attacking a particularly symbolic target and, instead, widen their target set and flaunt whatever works. Although the success of the Stuxnet worm suggests that individual targeting is possible, the attackers were not aiming for a precisely calibrated effect: The more damage, the better. Furthermore, the preparation for the attack was believed to have been years during which there

[20] Note that intelligence preparation of the cyberbattlefield, as it were, may differ sharply from everyday CNE. The former is concerned with understanding the target system well enough to understand what commands may make it act in a disruptive, destructive, or corrupted manner. It focuses on the instruction architecture of the target system. The latter tends to be a massive file-extraction exercise. It focuses on the content architecture of the target system.

المنارة للاستشارات

This content downloaded from
81.218.45.221 on Mon, 08 Nov 2021 13:47:06 UTC
All use subject to https://about.jstor.org/terms

was no reason for Natanz to suddenly increase its cybersecurity. By contrast, such preparation in the context of a war either presumes a very short war or risks stumbling when the security status of the targeted system shifts from a peacetime to a wartime mode.

The potential mismatch between effects and perceptions is another part of the same coin. The direct effects of a cyberattack may be obvious: The lights go out, for example. But, if the cyberattack is sufficiently complex, spreads very widely, or involves corrupted data (which, at first, appears valid), the true damage may be obscured even to the target. But perceptions rather than effects are the things to which the target state would react.

Last is the mismatch between perceptions and announcements whenever the damage is less than public. Obvious damage (such as the lights going out) is hard to misrepresent, particularly in our transparent times (once the damage is correctly characterized). But damage may not always be so obvious, especially if the system that is damaged does not have enough of a performance record to establish what normal operations look like. As noted, Iran's line on Stuxnet continued to evolve. Although announcements would seem secondary to perceptions, they may be the only information that third-party observers, the street, and even those outside the immediate circle of power will get.

Overall, the gearing between intent and consequence is multijointed and loose. Thus, a state may attempt escalation and (1) succeed, (2) fail but in such a way as to make no one the wiser, or (3) fail in ways that make it obvious that something was attempted but did not work. The latter simultaneously demonstrates malice and incompetence and may lead to overreaction as a way for the attacking state to regain the narrative. Alternatively, a state may just not respond when it could have, and something fails mysteriously anyway. It could be an accident, a rogue operative, a third-party state, or simply the inability of the target state to distinguish occasional failure from normal operations. So the target responds as if escalation had really taken place.

Finally, in cyberspace, the intent to react to escalation cannot necessarily be demonstrated as such. In the Vietnam War, escalation meant adding troops: easy to announce and verify. In cyberspace, neither the quality nor the number of the troops is obvious or can be reli-

المنارة للاستشارات

This content downloaded from
81.218.45.221 on Mon, 08 Nov 2021 13:47:06 UTC
All use subject to https://about.jstor.org/terms

ably monitored; indeed, these are usually highly secret. If there really is *any* ongoing conflict, there is no reason for a state *not* to assign all of its cyberattackers to the effort. Unlike, say, soldiers or sailors, they do not have to be deployed around the world in case another war breaks out, and it is not as if they cost more deployed than standing around.[21] The effects of making a greater effort may be long in coming; finding vulnerabilities is more like an investment, in which throwing more people at finding vulnerabilities produces more vulnerabilities—if they exist at all—only after a certain amount of time.[22] Furthermore, many of the best hacks are unnoticed by their victims until inexplicable failures start to mount. Only outputs count.

### Inadvertent Escalation

A tit-for-tat strategy may also lead to unintended consequences, particularly if the red lines on each side are unannounced or, if announced, not compatible.

Figure 4.2 illustrates what may occur in a local war in which both parties have thresholds but define them differently. The attacker, in this example, the United States, starts by hacking into the target's afloat naval supply facility database in order to scramble its contents. The target takes this as a cyberattack on military support and responds by hacking into the software system that controls Guam's port, to do likewise. The United States takes this to be an attack on the homeland (Guam being a U.S. territory), and it hacks into the software that controls port operations on the target's mainland. The target takes this as an attack on its civilian infrastructure. And so on.

All this escalation takes place even though neither side, at any time, believes that it is escalating. Each side is carrying out operations inside the boundaries within which the other side is already working. Yet, between the two, escalation happens. Although similar issues

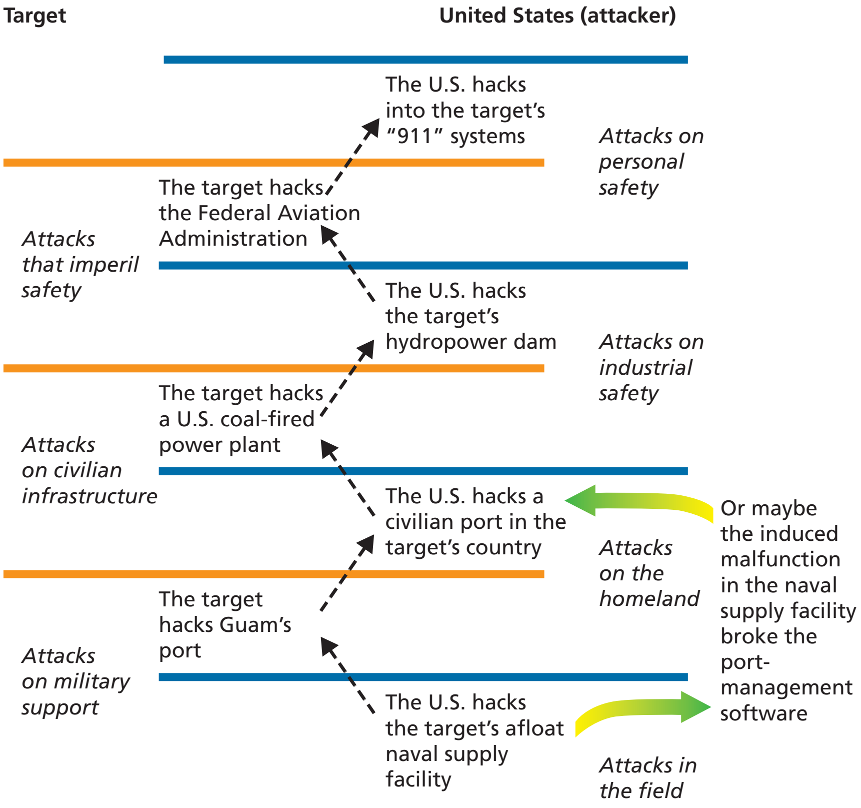
[21] The same claim cannot be made for cyberespionage, in which one hesitates to pull cyberwarriors from one country of interest, such as China, just because they may be useful to deal with a conflict elsewhere. But cyberattackers have no alternative cyberattack activity in peacetime.

[22] Stuxnet was estimated to have a gestation of a year, give or take a factor or two, and that may have been *after* the necessary zero-day attacks were discovered.

المنارة للاستشارات

This content downloaded from
81.218.45.221 on Mon, 08 Nov 2021 13:47:06 UTC
All use subject to https://about.jstor.org/terms

**Figure 4.2**
**An Inadvertent Path to Mutual Escalation**


Target
United States (attacker)
The U.S. hacks into the target's "911" systems
Attacks on personal safety
The target hacks the Federal Aviation Administration
Attacks that imperil safety
The U.S. hacks the target's hydropower dam
Attacks on industrial safety
The target hacks a U.S. coal-fired power plant
Attacks on civilian infrastructure
The U.S. hacks a civilian port in the target's country
Or maybe the induced malfunction in the naval supply facility broke the port-management software
Attacks on the homeland
The target hacks Guam's port
Attacks on military support
The U.S. hacks the target's afloat naval supply facility
Attacks in the field

RAND *MG1215-4.2*

bedevil escalation management in the kinetic world, the United States has no reasonable fear of having its homeland touched by another state in the physical world.[23] No such guarantees exist in cyberspace.

Alas, asymmetries between opponents will complicate tacit agreements on what to leave intact in the cyberworld, just as they do in the physical world. A local conflict between the United States and China over Taiwan will take place much closer to China: Agreeing

[23] Terrorism constitutes an exception, but one that is limited by virtue of the kind of weapons that can be brought into the United States and close to the target without being detected.

This content downloaded from
81.218.45.221 on Mon, 08 Nov 2021 13:47:06 UTC
All use subject to https://about.jstor.org/terms

that homeland ports are off-limits favors China because the gains to it from attacking embarking ports in, say, California are likely to be quite diffuse given the long steaming times. The reverse favors the United States. One country may use coal to generate its electricity; the other, hydropower. A policy that has each side refrain, for safety reasons, from interfering with dam controls unfairly penalizes the coal-using state; only its electrical generating capacity remains at risk. States that have built dedicated communication lines for defense are disadvantaged against states that must depend on dual-use infrastructures if both agree not to target dual-use nodes routers and switches. States that feed intelligence to "patriotic" hackers to carry out cyberattacks are at an advantage over those who depend on their own employees if the onus against cyberattacks is levied only against warfighters acting under state command.

Without announced red lines, states have to calculate how far they can go without touching the other side's nerves—and the extent to which adversaries will game such calculations. Similar issues associated with physical attacks can be dealt with through geographical limitations on combat: e.g., Northern Watch operations (circa 1993) did not extend below the 36th parallel. Boundaries in cyberspace are harder to define and confine. A reported U.S. strike on a jihadist web site supposedly took out 300 servers around the world.[24] Indeed, information system support for combat operations need not be anywhere near the conflict, RF bandwidth permitting; they are more survivable if they are not. So, a subtle adversary may deliberately outsource such processing to server clouds located in third-party countries. Thus, the useful boundaries have to be logical rather than physical ones. Unfortunately, as Schelling points out, such boundaries will limit the activities of both sides only if they are negotiated or obvious (e.g., stopping at the river's edge).[25] Otherwise, they seem arbitrary and meaningless, and therefore not credible guides to the other side's red lines; or, alternatively, concocted to favor the side that advocates them. The nuclear threshold was

[24] Ellen Nakashima, "Dismantling of Saudi-CIA Web Site Illustrates Need for Clearer Cyberwar Policies," *Washington Post*, March 19, 2010, p. A1.

[25] His theory of the focal point was developed in Schelling, 1960, pp. 53–80.

المنارة للاستشارات

This content downloaded from
81.218.45.221 on Mon, 08 Nov 2021 13:47:06 UTC
All use subject to https://about.jstor.org/terms

one such boundary. The distinction between fatal and nonfatal cyberattacks may be another. Avoiding the strategic path is a little trickier because a cyberattack can run this escalation path without the attacker and, for a while, even the target, realizing as much. Although the dual-use nature of some command, control, communications, computers, intelligence, surveillance, and reconnaissance systems may present similar difficulties for physical escalation, such problems are trickier in cyberspace to the extent that the virtual connections between systems are less visible. Thus, it is difficult to ascertain whether a strategic system was or was not depending on some capability or utility that was knocked out by a cyberattack meant to cripple a conventional capability. Indeed, if the wiring diagram between systems is sufficiently complicated, the target may not know that its strategic systems have been crippled until afterward.[26]

Finally, because the collateral effects of cyberwar are poorly understood, escalation-management strategies have to reflect the possibility of accidents. As shown on the right of Figure 4.2, the attack on the target's afloat naval supply facility may corrupt information, thus breaking the port-management software in that country (how was the United States to know that the target's port-management software did not do a sanity check on the information coming in from its ships?). Such accidents give further impetus to escalation in an environment in which both sides cannot bear escalation without matching it. Incidentally, no state should count on being able to argue that some effect was an accident, that it will not be repeated, and that accidents do not justify counterescalation by the other side. States rarely apologize even

[26] The strategic question of whether a state in a nonnuclear confrontation should raise a shadow over nuclear systems as part of its brinkmanship strategy is a separate issue not mooted here. If a state concludes against such a strategy, its policy on the use of cyberweapons should conform by staying as far away from the other side's nuclear C2 as it can (short of clear evidence that strategic weapons are about to fly or the threat to release them has already been made). Once a state thinks that its C2 is weak, it starts worrying about whether it has to use nuclear weapons while it still has control over them (whether such logic would apply if it fears that its systems might already be disrupted beforehand is a different issue). Similarly, if it starts to distrust its strategic surveillance, it may allow itself to make launch decisions based on less reliable but more trustworthy (for their not having been attacked) systems.

المنارة للاستشارات

This content downloaded from
81.218.45.221 on Mon, 08 Nov 2021 13:47:06 UTC
All use subject to https://about.jstor.org/terms

when wrong, and victimized states rarely settle for mere apology; reparations *during* wartime are even less common.[27] More typically, once a breach has been made, it tends to be exploited with vigor rather than backed away from.[28]

## Escalation into Kinetic Warfare

Under what circumstances can an attack limited to cyberspace or a conflict carried on by both sides only in cyberspace escalate into kinetic warfare and therefore violence? Are the two realms considered distinct and therefore unrelated, or are they part of the same continuum of force? Iran did not respond to Stuxnet with violence against the United States or Israel (but nor did Syria, for that matter, respond with violence to Israel's destruction of a purported nuclear reaction in 2007 despite hints that Israel used cyberwar techniques to help with air attack).

Several considerations merit note.

First, signals could be indicative. The more that a state has declared that it would respond to a cyberattack (that crossed some threshold), the greater the loss in face if it does not. If the attacker has few assets at risk from cyberwar (e.g., the Democratic People's Republic of Korea [DPRK]), the choice becomes one of either not responding meaningfully or responding with physical force. Likewise, the more that a state has rejected the idea of limiting a response to in kind, and the more it has embraced the concept of cross-domain deterrence (consider the smokestack reference earlier), the greater its odds of crossing from the virtual to the real world.

Second, hostile or at least bumptious action in certain domains seems to strike closer to home than do others. The United States and the Soviet Union had many incidents at sea, as noted earlier, and none

[27] The United States never apologized after downing an Iranian Airbus in 1988, although it did pay $62 million to settle subsequent claims eight years later. In 1904, the Imperial Russian fleet, thinking that it saw Japanese warships, attacked British fishermen and almost precipitated a war with England. See Gavin Weightman, *Industrial Revolutionaries: The Making of the Modern World 1776–1914*, London: Grove Atlantic, 2009, pp. 342–345.

[28] Kahn, 1965, p. 127.

المنارة للاستشارات

This content downloaded from
81.218.45.221 on Mon, 08 Nov 2021 13:47:06 UTC
All use subject to https://about.jstor.org/terms

of them escalated into actual war. The United States did not go to war when the DPRK captured the USS *Pueblo* in 1968 (nor break diplomatic relations with Israel over the sinking of the USS *Liberty* in 1967). Similarly, both sides ran active espionage operations against one another, and, with the possible exception of the furor associated with the Soviet downing of a U-2 aircraft, none of them seriously rippled the surface. This pattern has continued with espionage between Israel and its foes. Supposedly, U.S. and Soviet aircraft engaged one another during the Korea War without creating a broader crisis. By contrast, incidents involving Army soldiers (such as the death of MAJ Arthur D. Nicholson Jr. by East Germany in 1985, or the axe murder of 1LT Mark Barrett by North Koreans in 1976) seem to have had greater echoes. Would a cyberattack on the homeland be considered akin to a naval or intelligence incident and thus handled within its own channels? Or would it be considered akin to an army or homeland incident and thus lead to crisis and perhaps the use of force?

Third, the decision to use force—which is, in many cases, tantamount to starting a war—involves answering a set of questions, many of which have nothing to do with the incident that precipitated it. A rational state would ask itself, what would be gained by going to war? At what price? With what risk? But states, except for the most-obvious aggressors (e.g., Nazi Germany), often tell themselves that they are going to war because they have no choice other than to do so. That is, a world in which they do not go to war would be intolerable. Or the decision to refrain from war would only postpone conflict, not eliminate its possibility; when war came, outcomes would be worse, perhaps catastrophic.[29] Hence the question, what about a cyberattack would convince a state that it had no choice but to go to war? Clearly, the prospect of further cyberattacks would have to be considered intolerable. But the rationale to "take arms against a sea of troubles, and by opposing end them" presupposes that the use of force can end the pros-

[29] Examples may include a Wilhelmine Germany facing a steadily strengthening Russia and fearing encirclement, a Japan facing an economically devastating cutoff of raw materials, or an Iraq whose ability to pay war debts was being seriously crimped by Kuwaiti stubbornness about oil markets.

المنارة للاستشارات

This content downloaded from
81.218.45.221 on Mon, 08 Nov 2021 13:47:06 UTC
All use subject to https://about.jstor.org/terms

pect of cyberattacks. But can they? Given the difficulty of disarming hackers, such a prospect would appear to be dim. If the hackers capable of causing so much trouble emigrated, even occupation of their countries would not necessarily end their capabilities (although it would stoke revanchist motivations). That leaves, as a rationale for the use of force, the prospect of deterrence. A state punished severely enough for having launched cyberattacks against another might hesitate before doing it again; states that are watching may feel similarly disinclined. But this logic presumes that the state in question, as well as onlookers, convinces itself that it was the cyberattack that led to the use of force.

## Escalation into Economic Warfare

Another source of crisis exacerbation is the tendency for a trade war to overtake and become proxy for a budding cyberwar.

Indeed, it is hard to imagine any serious strategic cyberwar between two trade-linked states that does *not* become a trade war, and part of the art of managing a cybercrisis with a trading state is how to manage such fallout. This cuts two ways. A state can work to ensure that little of the cyberwar spills into the trading arena. Or it can use the threat of a trade war, coupled with the credible ability to wage one, to terminate a nascent cyberwar.

To illustrate, take the conflict scenario described in Chapter Six of *Cyber War*.[30] China starts by claiming all the South China Sea. The United States says no and conducts exercises with some newfound Asian friends. The United States leads the attack in cyberspace, first by sending China a warning in the form of an image of a sinking ship emailed from within China's supposedly closed military network, and then by turning off the power around the ports from which a potential Chinese amphibious invasion of disputed islands is being assembled—which unfortunately blacks out the entire province of Guangdong. This China considers escalatory. China retaliates in kind—and also

[30] Richard A. Clarke and Robert K. Knake, *Cyber War: The Next Threat to National Security and What to Do About It*, New York: Ecco, 2010, pp. 179–218.

المنارة للاستشارات

This content downloaded from
81.218.45.221 on Mon, 08 Nov 2021 13:47:06 UTC
All use subject to https://about.jstor.org/terms

blacks out more of the West Coast than intended. Things go south quickly: Key financial databases are scrambled, and the control computers for the major U.S. railroads and airlines go down. So the United States ups the ante, only to discover that China has disconnected itself from the Internet, thus blocking the most obvious route into China's cyberspace. It also phased down power interconnects among its regional power grids, thus limiting the possibility of cascading failures. Finally, China placed its railroads under manual control. In the end, the United States decides that it has less stomach for cyberconflict than the Chinese appear to and essentially throws in the towel, but not without first sending more carriers into the area.

Clarke's *Cyber War* scenario is decided by each side's susceptibility to a cyberwar, but the contest ends quickly before each side's susceptibility to a trade war is fully tested. Granted, a weekend (the interval over which the entire conflict takes place) hardly provides much time for a trade war. Even so, the economic ramifications of what was described in the scenario merit contemplation. China's willingness to cut itself from the Internet is likely to affect China's ability to export. China's export sector, much of which is products made to order for large customers, depends on large data flows of product specifications from U.S. manufacturers and on sales from U.S. marketers. The United States exports a lot less than it imports from China, and a large share of its exports are long-production items, such as aircraft, that, by virtue of long production runs, may be less sensitive than short-production items to temporary information outages. The asymmetry favors the United States. If the disruption lasts more than a few weeks, Western investors in China stand to lose a great deal of money. China's attraction as a manufacturing base would dip relative to other low-cost producers elsewhere in Asia and Latin America. China's physical investments in the West, although growing, are far smaller than the West's investments there.[31] A large share of China's investments outside China is in

[31] Chinese investments in the West have tended to be portfolio investments, such as stocks, bonds, and other financial instruments; until 2009, China's direct investments in the United States averaged $500 million per year or less, compared with several billion dollars per year coming from the United States, according to statistics from the U.S.-China Business Coun-

المنارة للاستشارات

This content downloaded from
81.218.45.221 on Mon, 08 Nov 2021 13:47:06 UTC
All use subject to https://about.jstor.org/terms

commodity extraction, in which the questions, such as the protection of intellectual property in cyberspace, are nugatory.

Even today's cyberespionage can be economically costly for China; the 2010 Google incident (in which the company's systems were penetrated and source code stolen) has reinforced the wariness that many Western corporations have felt at locating too much intellectual property where it can be stolen, and China cannot have been pleased when GE's president criticized China's challenging business conditions.[32]

Were China's cyberoffensive to include supply-chain attacks,[33] the damage to the United States may be sharper, depending on how many zombie components exist in U.S. systems and whether they can be accessed and activated by hackers when needed.[34] A scenario of a supply-chain attack *outside the context of war* is implausible largely because many zombie components can be replaced over time: admittedly, months and perhaps years compared with the hours and days associated with restoring penetrated systems. But, again, there would be a severe risk to China's export base—particularly its economic export base—that would follow the revelation of a deliberate supply-chain attack. The echoes may well last a generation.[35]

cil. See Thilo Hanemann, "Chinese FDI in the United States: Q4 2011 Update," Rhodium Group, April 4, 2012.

[32] Heidi N. Moore, "GE's Jeff Immelt Says It Out Loud About China," *CNNMoney*, July 2, 2010.

[33] Which, admittedly, the Chinese may have more grounds to be afraid of than the United States would, given the U.S. dominance in software.

[34] A zombie computer is a user computer that a hacker can also control.

[35] In 2010, following a row with Japan over disputed islands, the supply of rare-earth minerals from China was suddenly tightened. Because rare-earth minerals are predominantly used in the electronic sector, this may have been an attempt to pressure one of Japan's leading export industries. Rare-earth minerals, name aside, however, are not really rare. The United States used to mine them in the Mojave Desert and could resurrect such a capability if customers started to get nervous enough to pay a premium for alternative supplies. This incident raises questions of what trade-offs China is willing to make between fostering a reputation as a reliable supplier and using its manufacturing position to pursue national security goals (Keith Bradsher, "Amid Tension, China Blocks Vital Exports to Japan," *New York Times*, September 22, 2010b).

المنارة للاستشارات

This content downloaded from
81.218.45.221 on Mon, 08 Nov 2021 13:47:06 976 12:34:56 UTC
All use subject to https://about.jstor.org/terms

All this, incidentally, may take place without U.S. policy pushing a trade war, which would violate many trade agreements unless open hostilities were going on. It suffices that the disruption through which China is willing to put its own industry in order to make a political point shifts the calculus of thousands of independent decisionmakers outside China. Corporations may be reluctant to complain in public lest they alone face China's wrath, but there are many more-subtle ways of registering dissatisfaction, not least of which is by doing nothing. Suddenly, China notices that no one returns its calls anymore.

The decision to allow or even encourage a freelance response to a cyberattack has two sides. The argument for control is that it permits the United States, as a target, to manage the crisis through explicit or implicit negotiations. The argument against control is that an attacker is more likely to be deterred by the unpredictable reactions of thousands that cannot be individually coerced than by the well-considered actions of a state that can be. Individuals can be inhibited by the prospect that their country may suffer, but, unless they are worried about getting caught *and* the act is proscribed (which does not necessarily apply to, say, a refusal to invest in China or buy Chinese products),[36] they themselves are not at risk. The logic is similar for a state-encouraged response. States that can create sufficient doubt that they are behind the "people's" response may escape punishment for encouraging a vigilante response. That noted, this is a trick that is harder for a government operating in a transparent society to carry out.

A freelance cyberresponse may be more likely than a freelance trade war but less worrisome. States do not have a monopoly on clever hackers but, in most cases, cleverness alone does not suffice to carry out damaging cyberattacks; it takes intelligence on targets, notably on the processes that may go haywire if such information systems are attacked. It is one thing, for instance, to make a system stop working; such systems can often be repaired or their faults routed around in days. It is quite another to make it work in such a way as to mislead

[36] The larger the organization, the more likely it is to have a track record of investing in or buying from China and the more likely a sudden change in its investment and purchasing behavior is likely to be noticed.

المنارة للاستشارات

This content downloaded from
81.218.45.221 on Mon, 08 Nov 2021 13:47:06 UTC
All use subject to https://about.jstor.org/terms

decisionmaking, corrupt data irreparably, or interfere with some process control and wreak serious havoc. Even if they can produce one or two audacious attacks, clever hackers suddenly aroused to fury will generally not have such intelligence with which to work. Deep intelligence is the province of states.

In deciding whether to escalate from cyberwar to a trade war, several other factors enter the equation. Is the cost of a trade war low compared with whatever concessions are entailed in losing a cyberwar? Can a credible threat to do so convince the other side not to take advantage of its superior cyberwar capability? Will the other side back down first? How much damage would a trade war wreak on the world trade system? Would the ever-tightening chain of global sourcing make everyone, including noncombatant states, worse off? Will there be pressure to carry on from potential winners of a trade war, countries that want to sell to one side but no longer have to compete with imports from its rival? These are familiar questions to any strategist contemplating a contest in which two states can only hurt one another (as well as bystanders) and so the first one who cries "uncle" loses.

## Sub-Rosa Escalation

Another escalation option that might communicate displeasure to the other side without necessarily provoking it to respond is to strike systems whose malfunctioning will not be public even if apparent to the state's leadership.[37] The only entities that will *supposedly* know about the attack are the attacker, the target, and those to which either side confesses. Such limitations are meant to ease the pressure on the target to respond by escalation because no one can lose face (before the whole world) by backing off or not appearing tough enough. Sub-rosa options are generally unavailable to attackers in the physical world. Alternatively, cyberwarriors on both sides may be overcome by their own cleverness and create or exacerbate a crisis they cannot manage in

[37] This section and the accompanying Appendix B expand material that appeared in Libicki, 2009, pp. 128–129.

المنارة للاستشارات

This content downloaded from
81.218.45.221 on Mon, 08 Nov 2021 13:47:06 UTC
All use subject to https://about.jstor.org/terms

carrying out attacks and cyberattacks about which they thought only they knew. Done right, therefore, sub-rosa responses are likely to be less destabilizing than overt attacks and responses; done wrong, perhaps not so much. Appendix B examines sub-rosa operations within a broader three-by-three matrix of overt, obvious, and covert cyberattacks and responses.

Sub-rosa cyberattacks can be quite tempting, particularly among those within covert ranks. No one has to produce evidence of attribution. There is also less pressure to reveal the particulars (methodologies and targets) of the original attacks. Thus, the victims can pretend that nothing happened if they believe that they have no good counterescalation options or wish to contain the level of overall damage. Indeed, there are many reasons that carrying out covert operations in cyberspace is easier than in the physical world: e.g., fewer potential prisoners.

Unfortunately, what is most attractive to some becomes a weakness to others. Those who work in the highly classified arena can avoid the *public* oversight under which the more-overt parts of the national security community operate.[38] If the attacker wishes to justify its actions, it has more control over what evidence is collected and presented; it has less to fear from contradictory material provided by neutral or hostile parties. It avoids having to answer the question, if the evidence of who carried out the original attack will be unconvincing to others, how good can it really be? Members of the covert community, despite their personal probity and honesty, tend to operate in a sealed world. Mistakes can go uncorrected for longer than those made by overt operators. When actions are criticized, members of the covert community tend to circle the wagons. Even those who argue that members of *one's own* covert community are like everybody else, only in different professions, the same may not hold for members of *other states'* covert community, in which rule of law is generally and noticeably weaker.

The second problem with sub-rosa warfare is that each side's strategy is hostage to the discretion exercised by the other side, not to men-

[38] This is not to say they get less oversight, overall, but it is necessarily by those that have access to the same information compartments.

المنارة للاستشارات

This content downloaded from
81.218.45.221 on Mon, 08 Nov 2021 13:47:06 UTC
All use subject to https://about.jstor.org/terms

tion accident and error. Once revelations start, many parties will be embarrassed—not only the attackers on both sides but also the targets for allowing vulnerabilities to pervade their system and covering up after these vulnerabilities were exploited. Although a primary rationale for keeping matters covert is to facilitate later settlement, covert communities are not always motivated by the desire to reach accommodation with the other side. Covert communities, by their nature, distrust all other covert communities. So, each side has to weigh whether it is better off pulling back the shades on these sub-rosa exchanges or letting matters continue their subterranean course. The result may be a game of chicken. Each knows that revelation will make its side look bad not only to the public but perhaps also to its own masters, but each may hope that the *threat* of revelation may make the other side accede. Each side may therefore be in a position to concede things to hide its mutual activities in ways that might be impossible were its "negotiations" subject to public scrutiny.

Attacking covertly means not pursuing targets that serve the public (or private groups sufficiently large that having everyone remain silent is unlikely). Eligible targets are those that belong to parts of the government or to internal systems of institutions permitted and likely to keep matters private. Two ironies follow. One is that the best targets of sub-rosa cyberattacks are those whose workings are not only hidden but whose existence target states are reluctant to admit in the first place.[39] That noted, many such systems tend to be air-gapped and thus very hard targets. The other is that open societies, such as the United States, do not offer good targets for a sub-rosa attack because of the difficulty of keeping such attacks secret in such societies. Closed societies offer more good targets for sub-rosa attacks. Similarly, because

[39] In the wake of the controversy over the Defense Advanced Research Project Agency's Total Information Awareness program, funding was ended. If, as many believe, the program went underground into the intelligence community, those that run such systems may be quite reluctant to admit that they exist.

المنارة للاستشارات

This content downloaded from
81.218.45.221 on Mon, 08 Nov 2021 13:47:06 UTC
All use subject to https://about.jstor.org/terms

secrecy is emphasized in war, states at war offer more sub-rosa targets than those at peace.[40]

There is, incidentally, a world of difference between a deterrence strategy that assumes a public response and the option to go public. Threatening to go public with an act of escalation that may affect public opinion (e.g., by its audacity) *after the fact* is like relinquishing the steering wheel to an enraged public.[41] Once the crisis starts, however, the national security elite would be acting against type to relinquish that sort of control.[42]

## Managing the Third-Party Problem

Escalation-management strategies also have to contend with the problem of distinguishing attacks by third parties from those of the adversary. One saving grace is that the third-party problem is different in wartime. Against a background of full-bore cyberattacks, third parties have to make a larger splash than they do in peacetime to be considered escalation. In peacetime, a state that has been attacked in cyberspace and does nothing has to explain to its public and foreign observers why. In wartime, it can credibly argue that it is already doing all it can against the adversary and that a failure to escalate is not a display of cowardice. Because each side will naturally assume that its enemy on the battlefield is responsible for the escalatory cyberattack, the third parties do not have to strain to imitate the signature of a particular state. Even if there are doubts, the state that is the target of the third party is more likely to respond as if attacked by its battlefield foe if it reasons that doing so will not create a new enemy. Conversely, the possibility that escalation could have been carried out by third parties cre-

[40] There may be other bureaucratic reasons that the sub-rosa character of the operations remains. Inertia is one. The reluctance to declassify what were previously highly classified activities is another.

[41] Schelling, 1960.

[42] Leslie H. Gelb and Richard K. Betts, *The Irony of Vietnam: The System Worked*, Washington, D.C.: Brookings Institution, 1979.

المنارة للاستشارات

This content downloaded from
81.218.45.221 on Mon, 08 Nov 2021 13:47:06 UTC
All use subject to https://about.jstor.org/terms

ates an excuse for *not* counterescalating, even for an adversary's attack, until attribution is sorted out.

Third parties can create crises in wartime in ways unavailable in peacetime. A cyberattack on a strategic system (which should be nearly impossible but is conceivable) may be considered inexplicable in peacetime. A similar attack in wartime could be considered a precursor to escalation from conventional to strategic because such escalation is quite plausible. Fortunately, because the easy targets will have already been taken offline or hardened early in a war and the harder targets will require considerable preparation, early participation by third parties may be relatively ineffectual. Over time, however, serious third parties can contribute a larger percentage of the total mischief if they take the time to focus on the target system, deepening their understanding of it, and looking patiently for vulnerabilities.

If escalation management requires controlling third parties, two questions arise: First, how can states determine whether attacks came from third parties rather than adversaries? Second, and far trickier, how can states prevent their adversaries from mistaking third-party attacks for their own attacks, particularly escalatory ones?

Determining who carried out an attack—a third party or the battlefield foe—uses some of the same techniques such a question would require in peacetime. In wartime, an attacker's access is both worse and better: worse because there are fewer day-to-day contacts, and better because some of the entry points may come from proximity to military conflict (e.g., an enemy UAV transmitter/receiver penetrating the battlefield). Furthermore, because the adversary is likely to be carrying out a larger number of attacks in wartime, particularly on military forces, there should be a larger body of evidence from which to distinguish the adversary's modus operandi from those of third parties.[43] Defenders can choose to distinguish attacks by battlefield foes from others by reasoning that their foes have no interest in wasting

[43] If the third party is attacking precisely to create further mischief between adversaries, what prevents it from copying one side's modus operandi as part of the ruse? The answer may be stated as a question: Can one copy a well-known modus operandi (which, having been used, has already set defenses against it in motion) and still carry out a successful attack?

المنارة للاستشارات

This content downloaded from
81.218.45.221 on Mon, 08 Nov 2021 13:47:06 UTC
All use subject to https://about.jstor.org/terms

their assets, notably their knowledge of the opponents' vulnerabilities, on low-impact attacks; thus, low-impact attacks are likely to have been carried out by others.

It is not easy to keep third-party attacks from inducing an adversary reaction. Warring parties rarely overflow with mutual trust. Having each side monitor the other's cyberwarriors to ensure that their attacks are limited in scope is not possible for an activity that requires deception to work.

If dealing with a foe that is less sophisticated and likely to overreact to cyberattacks against sensitive systems, one could monitor and immunize their systems against the attacks of others—that is, firewall such systems to ensure that no third parties get through. This may sound far-fetched, but some forms of the malware that convert systems into bots make it difficult for third parties to insert their own malware into such systems. If that is unappealing, a state can at least tell adversaries that some of their sensitive systems (that it does not intend to attack) have vulnerabilities so that they attend to such vulnerabilities before third parties exploit them. However, *finding* such vulnerabilities would require spying on such systems, which may itself raise suspicions.

## The Need for a Clean Shot

The problems of intrawar deterrence are similar in many respects to those of interwar deterrence, insofar as the *threat* to retaliate will work only if the adversary

- believes that it will be blamed[44]
- believes that the target has the means to carry out the deterrent threat

[44] This is a much smaller problem for intrawar deterrence because the usual reason to not respond in peacetime is the fear of starting a war—but, if the war has already started, such a fear is limited to the fear of the other side escalating.

المنارة للاستشارات

This content downloaded from
81.218.45.221 on Mon, 08 Nov 2021 13:47:06 UTC
All use subject to https://about.jstor.org/terms

- believes that the target has the will to carry out the deterrent threat even if it threatens to counter the target's reprisals with reprisals of its own
- believes that, if it does *not* cross a red line, it will not face escalation
- feels that its escalation has no compelling rationale that persuades it that it is militarily better off having escalated even after taking the target's potential response into account[45]
- does not fear losing too much face by complying (which argues for making such a threat implicitly or covertly)
- believes that the red lines are well-defined, straightforward to monitor, and considered fair—rather than one-sided, arbitrary, unfounded in customary law, or self-serving.

This is clearly a list of nontrivial length and content. Just as clearly, the success of a tit-for-tat strategy of intrawar deterrence has everything to do with what the other side believes. Thus, those that would adopt such a strategy have to have a fairly good read of the other side.

The problems do not end there if a state declares or strongly implies a tit-for-tat strategy and has defined red lines, and it is attacked anyway. It will have to either respond or give a good show of why it did not. It can claim that what the other side did was not escalatory by, somewhat incredibly, pretending that it did not cross a threshold or that it is unsure who did what and hope that the adversary does not take credit.[46] If that claim is unconvincing, the state may have to either

[45] This better-off logic does not apply in peacetime because a stand-alone cyberwar, incapable of destroying very much for very long, ends up becoming a battle of pain-making and pain tolerance and hence tends toward the mutually unsatisfactory when both sides weigh in. Thus, the prospects for peacetime deterrence, as problematic as it is, at least has some of the calculus in its favor compared with intrawar cyberdeterrence, in which mutual escalation can actually leave one side better off on the battlefield.

The target's potential response is particularly important if the worst possible reaction in cyberspace is a tolerable price to pay. This is no guarantee, however, that the target will not respond violently, if it can.

[46] That is, if one side wants to avoid having to respond to escalation by pretending that it was a third party that carried out an attack that crossed a red line (or would have crossed a red

المنارة للاستشارات

This content downloaded from
81.218.45.221 on Mon, 08 Nov 2021 13:47:06 UTC UTC
All use subject to https://about.jstor.org/terms

escalate when prudence would dictate otherwise or do nothing and lose credibility.

## Inference and Narrative

Escalation, by definition, is doing something different today from what one has done before. It leads to speculation about whether the adversary's intentions have changed or are different from once thought. Similarly, a state's failure to respond to escalation also gives rise to speculation about its attitudes.

Consider how a state's response to its adversary's cyberescalation may be read. What might others infer from a state's responding to cyberescalation with escalation of its own?

- The attack was detected and attributed correctly—a nontrivial achievement. A corruption attack or a destruction attack against a little-used but nevertheless critical function, such as backup, may go undetected.
- The state *can* escalate—also nontrivial. It means that the state has the technical know-how to breach barriers associated with targets that were previously untouched.
- The state *would* escalate. The state is not afraid of escalation; it cannot be cowed. Additionally, whatever inhibitions it had against hitting a class of targets no longer exists.
- The initial attack hurt or embarrassed the target state enough to convince it to carry out cyberattacks of the sort that it previously did not want or need to do. Or the discomfort was so great that the target state would escalate to really painful points in order to create a clear deterrent against carrying out such attacks.

line had its adversary carried it out), its strategy would be frustrated if the adversary stood up and claimed, "I did it!" That would put pressure on the target to respond.

المنارة للاستشارات

This content downloaded from
81.218.45.221 on Mon, 08 Nov 2021 13:47:06 UTC
All use subject to https://about.jstor.org/terms

- The state does not like risking casualties by responding kinetically, so it responds only in cyberspace.[47]
- The state is cruel and vicious, particularly if the response crosses red lines the adversary had yet to breach. It therefore must be heeded or, alternatively, destroyed.

The first three responses flatter the responding state as, respectively, adept on forensics, capable on offense, and steadfast. The last three point to a state that is, respectively, oversensitive, cowardly, and thuggish.

All this assumes that the retaliating state was, in fact, responding to an attack by the entity against which it retaliated. If no such attack took place, the state may be viewed as twitchy, trigger-happy, and ultimately incompetent. If an attack took place but from another entity, the state's confidence in its own attribution systems would be deemed misplaced. Or the attacker may convince itself that the retaliating state is dishonest about why it escalated and was just looking for an excuse (particularly if no such precipitating and escalatory attack took place).

Correspondingly, a state that failed to respond may allow the reverse implications to be drawn. That is, the state could not detect the attack, could detect the attack but was unsure who did it, or could not respond successfully. The state may have been cowed into not responding. Alternatively, it would not breach its ethical norms to respond, or it could afford to let such an attack pass.

The broad narrative that a state has used to frame its cyberspace policy may color its response options. A narrative that assumes bad things in cyberspace largely because systems are complex and fragile buys a state some time to consider its options after an attack. A corollary narrative that focuses on the faults of the defense rather than the fiendishness of the offense also makes it easier to avoid counterescalation.

The target state could make it appear that it retaliated when it did not. It could announce that its hackers have been unleashed (pre-

[47] Suppose that X attacks Y. Y responds but only in cyberspace. X infers that Y is a coward when it comes to violence, but X's inference is unfounded if Y just did not think that the damage from the attack rose to the level of justifying violence.

المنارة للاستشارات

This content downloaded from
81.218.45.221 on Mon, 08 Nov 2021 13:47:06 UTC
All use subject to https://about.jstor.org/terms

sumably, hackers had been leashed earlier). Faking a kinetic attack is very difficult, but faking a cyberattack is not because nearly everything about it is hidden. Such a stratagem would be the opposite of a sub-rosa response, and the claimed retaliatory attack would have to have nonobvious effects (e.g., corruption rather than disruption). An opponent that believes this may well divert resources to calculating which of its information stores or algorithms have been tampered with. What it concludes if it finds something suspicious—for any number of other reasons—is another issue.

Should a state, then, escalate based on what its opponents succeed in doing or on what they tried to do? If the purpose of escalation management is to inhibit what foes try to do rather than what they succeed in doing, then attempts alone suffice for a response. Yet, successful attacks illuminate intent much better than failed ones do: An armed man caught entering a building may clearly have been up to no good, but who was his target, and was his intent murder, assault, or intimidation? Furthermore, not only is the public unlikely to know of failed attempts, but, in some circumstances, the foe may be unsure why the attempt failed and thus may not be sure that the attempt registered with the target. So, the target loses less face when not responding to a failed attack.

States inclined toward retaliation may need to explain why particular targets that are out of bounds for kinetic attack are fair game for cyberattack (e.g., when is an attack on a port that supports operations in an offshore theater island *not* prefatory to an invasion of the adversary's homeland?).[48] The next question is obvious: Why would the victim state believe such a state, particularly if the attacking state

[48] Although a state could announce that it is eyeing a particular target system in order to elicit from the target any reason that such an attack should not be carried out, will the target use such warning to bulwark or isolate such systems or scream very loudly in the hopes that it can be spared even though nothing particularly critical was at stake? Would the target even find such a request legitimate? See, e.g., Lincoln P. Bloomfield, Jr., "National Security Fundamentals in the Space and Cyber Domains," *High Frontier*, Vol. 7, No. 1, November 2010, pp. 34–38.

There is no assurance that clear messaging at the leadership level between the United States and the adversary would serve as a brake on escalation in such a situation; but the absence of such communication would leave each side with no incentive or excuse for restraint.

المنارة للاستشارات

This content downloaded from
81.218.45.221 on Mon, 08 Nov 2021 13:47:06 UTC
All use subject to https://about.jstor.org/terms

suspects that the *only* purpose of any announcement would be to gain some military or strategic advantage? One answer may be that ancillary actions prefatory to a general escalation are absent. But that presupposes that the target of retaliation *can detect* such ancillary actions well enough to know that the cyberattack had a limited purpose, when a key purpose of any cyberattack is to persuade the adversary to doubt its information. So, whereas the problem of explaining escalation is not unique to cyberwar,[49] the use of cyberwar makes all explanations all the more suspect.

Inferences are even harder to draw when states are not unitary actors.[50] One bureaucratic faction in a warring state may carry out cyberattacks to rally the state's population behind its particular bent, say, in favor of greater belligerence, or against its particular *bête noire* (to take on country A when others want to take on country B).[51] Although kinetic attacks, particularly the larger ones, can be traced back, the source of a cyberattack may remain mysterious for a long time. Even leaders who seek calm can be frustrated by the difficulty of enforcing their writ on their minions—and, because a capacity for cyberwar needs only hackers, sufficiently detailed intelligence on the target, and a modicum of hardware, factions may have the requisite power to create considerable mischief. Retaliation by the target may well play

[49] As argued in Walter B. Slocombe, "Preplanned Operations," in Ashton B. Carter, John D. Steinbruner, and Charles A. Zraket, eds., *Managing Nuclear Operations*, Washington, D.C.: Brookings Institution, 1987, pp. 121–141, "How do you convince the other side that one's limited attacks are, in fact, limited?"

[50] Is China, for instance, a unified actor?

> [China's President Hu Jintao's] strange encounter with Defense Secretary Robert M. Gates here last week—in which [Hu] was apparently unaware that his own air force had just test-flown China's first stealth fighter—was only the latest case suggesting that he has been boxed in or circumvented by rival power centers. (David E. Sanger and Michael Wines, "China Leader's Limits Come into Focus as U.S. Visit Nears," *New York Times*, January 16, 2011)

[51] Japan's army circa 1941 was more interested in combat with China and perhaps Russia, while Japan's navy had its eye on the West's colonies in South and Southeast Asia and thus was itching to go after the UK and the United States. The United States in the 1790s found itself divided between factions that favored France and those that favored its wartime enemy, Britain.

المنارة للاستشارات

This content downloaded from
81.218.45.221 on Mon, 08 Nov 2021 13:47:06 UTC
All use subject to https://about.jstor.org/terms

into the hands of the aggressive rather than the more cautious faction; the former can display it as proof of the target's hostile nature. In such circumstances, the target state must ask, would the positive deterrence effect from counterescalation trump the negative effect from confirming the narrative of the more aggressive faction? If not, the target state may prefer to let the incident pass.

## Command and Control

C2 arrangements color escalation management in all forms of combat, but nowhere more so than in cyberspace. The problem arises with both the commanders and those they command.

### Commanders

Will commanders act appropriate to the crises, follow standard operating procedures, or enhance institutional interests? As Barry Posen, for instance, observed, "During the Cuban Missile Crisis, the U.S. Navy ran its blockade according to its traditional methods, disregarding President Kennedy's instructions," adding "orders to cease U-2 flights near the Soviet border were either not received, or were ignored; Soviet detection of these flights hindered the negotiations to end this crisis."[52] George Smoke argued that one of the reasons that Britain found itself mired in the Crimean War was that it perceived that Russia's devastating defeat of the Turkish fleet at Sinope was meant as an insult to the British themselves.[53] Britain implied that it would not respond if the Russians fought at sea as long as they did not attack a Turkish port. The czar concluded from this that naval action was acceptable as long as it took place at sea. But Russian admirals interpreted matters consistently with their desires and carried out their actions within the port of Sinope (without actually attacking the port facilities themselves). All that noted, cyberoperations do not have the long history of naval

[52] Barry R. Posen, "Inadvertent Nuclear War? Escalation and NATO's Northern Flank," *International Security*, Vol. 7, No. 2, Autumn 1982, pp. 28–54, pp. 32, 34.

[53] Smoke, 1997, p. 182.

المنارة للاستشارات

This content downloaded from
81.218.45.221 on Mon, 08 Nov 2021 13:47:06 UTC
All use subject to https://about.jstor.org/terms

operations. Whatever standard operating procedures exist are yet to be validated in a war or crisis against a competent enemy.

The influence of the cyberwarrior community's drive for status and recognition may play a large role. Like the U.S. Army Air Forces in the 1940s, cyberwarriors may wish to be seen as part of a military organization capable of creating strategic effects rather than just supporting other warfighters.[54] In a war, would they see more value in using their limited set of exploits against strategic targets? Would they disdain operations against military targets (that normally present little escalation risk if they can be engaged by kinetic means) in favor of strategic operations that carry a nontrivial risk of escalation?

The role of cyberwarriors within the regional combatant commands (COCOMs) colors the question for the United States. If, for instance, cyberwarriors were organized as teams reporting to a warfighting organization, such as an army division or even a regional combatant commander, then the subordination of community prerogatives to the total fight would more likely follow. In the United States, however, regional commands do not "own" cyberwarriors. All cyberoperations come under the C2 of USCYBERCOM, whose units are *not* chopped to combatant commanders but exercised directly.[55] USCYBERCOM reports to USSTRATCOM, whose other primary mission is nuclear and space forces. As LTG Keith Alexander emphasized,

> The Commander of U.S. Cyber Command will have freedom of action to conduct military operations in cyberspace based upon the authorities provided by the President, the Secretary of Defense, and the Commander, U.S. Strategic Command. Because cyberspace is not generally bounded by geography, the Commander of U.S. Cyber Command will have to *coordinate* with U.S. agencies

[54] Cyberwarriors have not pressed to become their own corps, much less their own service. Although we put forth the case for a separate information corps nearly 20 years ago, the purpose of such a corps was to generate a joint picture of the battlefield based on the coordinated operation and analysis of sensors, not to carry out information warfare (as it was then called). See Martin C. Libicki and James A. Hazlett, "Do We Need an Information Corps?" *Joint Forces Quarterly*, Vol. 2, Autumn 1993, pp. 88–97.

[55] *Chopping* a unit means to allow a unit under one's command to work for another commander for the time being.

المنارة للاستشارات

This content downloaded from
81.218.45.221 on Mon, 08 Nov 2021 13:47:06 UTC
All use subject to https://about.jstor.org/terms

> and Combatant Commanders that would be affected by actions taken in cyberspace.[56]

Hence the question, current authorities notwithstanding, who ought to determine what cyberoperations are carried out during a military crisis or war? Institutional factors cannot be ignored.[57] In a war or military crisis, only USCYBERCOM will really know whether worthwhile strategic targets have vulnerabilities that can be exploited and to what effect, whereas the existence of kinetic targets is easier to demonstrate (e.g., by imagery). Insider knowledge may influence the options that the U.S. cyber commander presents to the regional COCOMs, if the latter get to select at all. Otherwise, a U.S. cyber commander may well select targets and take risks (or avoid taking them) that are inconsistent with how regional commanders would fight, if indeed the latter understand the risks and rewards of operations in cyberspace with sufficient detail at all. Yet it is the regional commanders who are the more knowledgeable about the most important factor in escalation management, the other side: what it thinks, what it infers about the U.S. posture, and where its red lines are drawn.[58] If USSTRATCOM backs up the cyber commander's subordinate commander (even that is unnecessary if the U.S. cyber commander becomes a combatant commander in his or her own right) when there are disagreements over targeting and operational procedures, the Secretary of Defense (SecDef) would have to arbitrate. Under such circumstances, regional commanders may not wish cyber to be the issue that gets raised to that level, particularly because the U.S. cyber commander will be the sole source of the details required to resolve the balance of risk and reward. Furthermore, requiring SecDef intervention will almost certainly slow the pace of battle.

[56] Alexander, 2010, p. 14. Emphasis added.

[57] See also U.S. Government Accountability Office, *Defense Department Cyber Efforts: More Detailed Guidance Needed to Ensure Military Services Develop Appropriate Cyberspace Capabilities*, Washington, D.C., GAO-11-421, May 2011.

[58] By contrast, the close relationship between USCYBERCOM and the National Security Agency (NSA) reduces the odds that intelligence gain/loss considerations will be ignored when attacks on targets threaten to reveal penetrations to the target, the fixing of which jeopardizes taps into systems that produce intelligence.

المنارة للاستشارات

This content downloaded from
81.218.45.221 on Mon, 08 Nov 2021 13:47:06 UTC
All use subject to https://about.jstor.org/terms

Escalation management also has to account for the power of the cyberwarrior community to force action when inaction may be called for. Consider that holes, once they are revealed to the defender, tend to be closed quickly. A cyberwarrior, on a limited mission, could exploit a vulnerability, discover that its exploitation can have vast if not necessarily precisely scoped effects, and beg for the authority to pursue action lest the vulnerability close forever. Active defense—in the sense of prompt action against machines on the attack—also presents opportunities for light-speed reaction that could lead to escalation that a little contemplation can foresee and forestall. The problem, of course, is worse, if the cyberwarriors are deliberately heedless of bounds on their actions, an issue covered next.

All this argues for two propositions. First, combatant commanders should have full control over cyberoperations whether or not they are deemed operational or strategic, if for no other reason so that their escalatory effects can be factored into the overall campaign plan.[59] Second, it may be useful for the United States to keep its cyberwar community under commands whose primary mission is the application of kinetic force, the better to remind everyone that cyberoperations exist to further the political ends—which, as Carl von Clausewitz observed, are the justifications for kinetic operations as well.[60]

## Those They Command

States that would manage escalation in cyberspace must have appropriate C2 of their cyberwarriors. Instructions on what to avoid must be clear, and the controls must be in place to ensure that such instructions are followed.

In the physical world, both command and control are getting better thanks to ever-more-ubiquitous surveillance and the proliferation of communication networks. The effects of war can be meticu-

[59] This does not imply that the regional commander would have access to all the tools possessed by USCYBERCOM because it may be advantageous to hold some tricks in reserve so they can be available for their greatest need, which may not necessarily be in the theater of operations at the time.

[60] Carl von Clausewitz, *On War*, Princeton, N.J.: Princeton University Press, 1989.

المنارة للاستشارات

This content downloaded from
81.218.45.221 on Mon, 08 Nov 2021 13:47:06 UTC
All use subject to https://about.jstor.org/terms

lously documented and attributed.[61] As more military equipment becomes digitized and thus capable of hosting copious log files, the prospect of knowing exactly who did what and when draws closer.

Not so in the cyberworld, in which keystrokes can come from anywhere. Standard operating procedures are a poor guide when one cannot say a priori exactly what the means of attack are, much less what the likely effects of attacks are. Any policy designed to attack up to some boundary but no further is subject to the two aforementioned differences: between intent and effect and between effect and perception. If one would act, clear and *thick* (to account for misunderstandings) margins of some sort have to be established.

The burden of margin-setting will differ depending on whether one is worried about careful, careless, or rogue cyberwarriors.

Careful cyberwarriors are those that pay as much attention to constraints as they do to results. For them, clarity is the goal. The constraints on their behavior could include how to attack and what results are wanted and unwanted under which circumstances. The bounds should be explicit, advertised, and stable against arbitrary change. The rules that say what actions are permissible in what situations should be codified in advance of crisis because, when the fighting starts, purposes are more fluid and not necessarily obvious to all. To make constraints work, it may be necessary to teach the basic principles of cyberwar as they apply to national security. Beyond such guidelines, however, the rules on how to attack or what constitutes nonexcessive damage may be too context-specific to be specific in advance.

Careless cyberwarriors mean to follow the rules but, in the heat of combat, may convince themselves that carrying out a clear operational mission trumps conformance with inevitably ambiguous guidelines. All the rules for careful cyberwarriors apply to careless ones, and the two may be indistinguishable. The application may vary: The actions of careless warriors are likely to drift over the borders, and, being human,

[61] Martin C. Libicki, David C. Gompert, David R. Frelinger, and Raymond Smith, *Byting Back—Regaining Information Superiority Against 21st-Century Insurgents: RAND Counterinsurgency Study—Volume 1*, Santa Monica, Calif.: RAND Corporation, MG-595/1-OSD, 2007, Chapter Four.

المنارة للاستشارات

This content downloaded from
81.218.45.221 on Mon, 08 Nov 2021 13:47:06 UTC
All use subject to https://about.jstor.org/terms

such warriors are likely to blame their trespasses on unclear guidance, the ambiguities of cyberspace, and even the target's behavior (e.g., turning off the electric power substation to disable government bureaus was not supposed to put hospital patients at risk; where were the latter's backup generators?). If careless cyberwarriors are a problem, one approach would be to limit the amount of intelligence with which *all* cyberwarriors are provided (e.g., avoid probing systems that will never be targets). But, given a wide enough range of contexts, what systems can one aver will *never* be targets?

Rogue warriors are those so eager to strike the target that they take their work home with them, sometimes literally. Trained and filled with intelligence at work, they carry out attacks from platforms or intermediate conduits that are very difficult to trace and out of sight of their supervisors. Rogue warriors will not respond to constraints when freelancing except as warnings about what to avoid appearing to do. Because they do not have to work in military formations or with unique military hardware, their operations are harder to detect and hence control than their equivalents in physical combat (e.g., the militias of developing nations). Not even keeping them chained to their desks in a military crisis will eliminate mischief if they have found how to contact their own bots from their desktop—although such behavior may be suppressed if they have to account for every keystroke. Effective militaries have ways of filtering out most such rogue warriors and engineering social controls that keep potential rogue warriors in the force from straying. Having done what they can, states then have to determine whether the risks of violating self-imposed constraints merit reducing every cyberwarrior's access to the intelligence and tools necessary to mount the more-sophisticated attacks.

المنارة للاستشارات

This content downloaded from
81.218.45.221 on Mon, 08 Nov 2021 13:47:06 UTC
All use subject to https://about.jstor.org/terms

## Conclusions

A state that would limit wartime cyberattacks against its society and out-of-theater military must pay attention to cyberescalation.[62] Avoiding escalation may be simpler if a war's goals are limited and actions follow accordingly. But fine-grained escalation management in cyberspace will remain tricky because of the difficult coupling between intentions, effects, and perceptions.

Escalation in cyberwar—particularly if cyber against cyber—is likely to be jerky rather than smooth. The kind of escalation presented by Herman Kahn, in which both sides feel their way up the proverbial escalation ladder to see who breaks first, is unlikely to characterize cyberwar (whether it characterizes any war is a separate question). What looks like a carefully calibrated ladder may, in practice, end up as a hodgepodge of sticky and bouncy rungs.[63] Thus, although Figure 4.2 shows a multistep ladder in the absence of well-defined and agreed-upon thresholds, a few large moves are more likely. Perhaps there will be only one escalation phase—from the unproblematic use of cyberattacks in an operational context against military targets, to an entirely problematic set of attacks that have or appear to have a strategic and coercive rationale against civilian targets.[64] Alternatively, the only attacks that may be deemed seriously escalatory are those that cross the border between instrumental (tactical) and general (strategic) or from legitimate to illegitimate. That, in turn, presupposes norms of what is one and what is the other, and such norms do not exist now and may not exist anytime soon.

Unlike other forms of warfare, the first use of a serious cyberattack could easily make states realize that the security-versus-convenience

[62] Because cyberattacks may lead to kinetic escalation, the importance of escalation management is not limited to the virtual realm.

[63] Sticky rungs are those from which one cannot rise; bouncy rungs are those from which one rises much farther than anticipated.

[64] Imagine a scenario in which the regional combatant commander makes an urgent request that a SAM site be knocked out. The cyber commander sees no way to get into the SAM site but knows that a "small" attack on the local power supply may have the same effect. The other side finds this "small" attack escalatory. And so on.

المنارة للاستشارات

This content downloaded from
81.218.45.221 on Mon, 08 Nov 2021 13:47:06 UTC
All use subject to https://about.jstor.org/terms

trade-off had tilted too far to convenience;[65] they will thus harden themselves quickly, making future cyberattacks more difficult. A cyberattacker that understands as much will necessarily want to frontload attacks knowing that attacks postponed are attacks denied. Furthermore, the *perceived* effects of cyberattacks tend to be more unpredictable than the effects of kinetic attacks.[66]

Proxy conflicts are particularly hazardous from the perspective of controlling crises by keeping matters in theater. The many potential third parties each have their own agendas, and physical boundaries are a relatively poor delineation of what is or is not a legitimate target. Nevertheless, both sides could take caution to isolate systems they put into theater from home systems, and each should remember that accidents happen. Wisdom also suggests postponing action against third parties, however annoying they may be.

States should also take the time to consider escalation carefully. There is little to be gained from an instant response. Cyberattacks cannot disarm another side's ability to respond in kind. True, cyberattacks cannot be frozen to be thawed out when needed; maintenance requires recurrent surveillance. But the timing of a response ought to be predicated on one's warfighting strategy, not a desire for speed-of-light responses.

Each state should also understand the other side's reaction to cyberescalation, notably what ethical norms it associates with cyberattacks vis-à-vis kinetic attacks and what others may infer about the attacker's intentions from such attacks. In cyberspace, as in other realms of warfare, "the defender frequently does not understand how threatening his behavior, though defensively motivated, may seem to the other side."[67]

[65] So, is the security-versus-convenience trade-off ipso facto tilted away from security? Perhaps necessarily; perhaps everyone has it right and no devastating cyberattacks are in the offing. More likely, some have tilted one way and some the other, and a major incident will excite the security laggards disproportionately.

[66] That noted, the psychological impact of the perceived effect has large random error terms for both kinetic and cyberwarfare.

[67] Posen, 1982, p. 33.

المنارة للاستشارات

This content downloaded from
81.218.45.221 on Mon, 08 Nov 2021 13:47:06 UTC
All use subject to https://about.jstor.org/terms

المنارة للاستشارات

This content downloaded from
81.218.45.221 on Mon, 08 Nov 2021 13:47:06 UTC
All use subject to https://about.jstor.org/terms