JOURNAL OF CYBERSECURITY

*Journal of Cybersecurity*, 3(1), 2017, 7–17
doi: 10.1093/cybsec/tyw015
Research paper

Research paper

# The role of effects, saliencies and norms in US Cyberwar doctrine

**Henry Farrell and Charles L. Glaser***

George Washington University

*Corresponding author: Email: cglaser@gwu.edu

Received 13 December 2016; accepted 14 December 2016

## Abstract

The US approach to cybersecurity implicitly rests on an effects-based logic. That is, it presumes that the key question determining how the US and others will respond to attacks is what effects they have. Whether the effects come about as a result of cyber means or kinetic means is largely irrelevant. In this article, we explore this logic further, focusing on the question of when the US should deploy cyber responses and when kinetic. We find that under a simple effects-based logic, kinetic responses will often be more effective than cyber responses, although we explain that cyber attacks that 'leave something to chance' may be an effective deterrent under some circumstances. We next develop a richer understandings of actors' expectations by employing the concepts of focal points and saliencies. In this framework, kinetic responses may be considered too escalatory, and therefore less attractive under many circumstances. If there are 'focal points' emerging, under which cyber attacks are seen as qualitatively distinct from kinetic attacks, then crossing a saliency may appear escalatory, even if the actual effects of the kinetic and cyber attackes are identifical. Finally, we examine nascent norms around cyber, suggesting that the US may wish to consider promoting a norm against large scale attacks on civilian infrastructure, and evaluating the prospects for a norm against cyber attacks on nuclear command and control systems.

**Key words**: cyber; nuclear; deterrence; focal point; salience

How should the USA respond if an adversary employs cyberattacks to damage the US homeland or weaken its military capabilities? Closely related, what threats should the USA issue to deter these attacks? The most obvious answer may be that cyberattacks should be met with cyber retaliation. Careful examination of these questions shows, however, that under a variety of conditions the USA should retaliate with conventional military attacks—that is, kinetic attacks. On the flipside, are there situations in which the USA should employ cyberattacks to improve its prospects for success in a conventional war?

To analyze these questions, we draw upon and combine three approaches—effects, saliencies and norms. We begin with a basic effects-based logic—that is, decisions about deterrence and warfighting should be based on the effect a US attack will have, not on the means via which that effect is produced; if kinetic retaliation and cyber retaliation would inflict comparable costs, then there is no obvious reason to favor one over the other. We then draw upon the concepts of focal points and saliences to add useful distinctions. This is necessary because the pure effects-based logic is likely too sparse—states may perceive different forms of retaliation that do equal damage (that is, are equally costly) as differently punishing and differently escalatory. Finally, we consider the possibility that norms against certain types of cyberattacks should impose limits on US cyber doctrine. Although such norms have not yet been established, beyond those that apply generally to the laws of war, we discuss a couple of possibilities, as well as the barriers to their achievement.

Current US cyber doctrine is consistent with a basic effects-based approach, making clear that the US envisions the possibility of kinetic attacks in response to cyberattacks: 'The United States will continue to respond to cyberattacks against U.S. interests at a time, in a manner, and in a place of our choosing, using appropriate

instruments of U.S. power.' On the flipside, the strategy also suggests that the USA might rely on cyberattacks to contribute to US efforts that have not yet involved cyberattacks: 'the United States military might use cyber operations to terminate an ongoing conflict on U.S. terms, or to disrupt an adversary's military systems to prevent the use of force against U.S. interests'. The strategy emphasizes that cyber capabilities will be integrated with the full range of other US fighting capabilities:

> DoD should be able to use cyber operations to disrupt an adversary's command and control networks, military-related critical infrastructure, and weapons capabilities … .To ensure unity of effort, DoD will enable combatant commands to plan and synchronize cyber operations with kinetic operations across all domains of military operations (11, 5, 14) [1].

Less clear is whether US doctrine goes beyond this effects-based approach to include saliences and norms (except for the laws of armed conflict (6) [1]).

Influential analyses of cyber strategy question a purely effects-based approach, worrying that a kinetic response to a cyberattack could constitute a dangerous escalation. The alternative perspective argues that other considerations—beyond the amount of damage that an attack would inflict—may influence an adversary's understanding of and reaction to an attack. For example, Herbert Lin cautions that '[n]ations involved in a cyber-only conflict may have an interest in refraining from a kinetic response—for example, they may believe kinetic operations would be too provocative and might result in an undesired escalation of the conflict' (65) [2]. Martin Libicki offers a similarly cautious perspective; while not ruling out kinetic responses, he argues that a kinetic response 'would trade the limited risks of cyberescalation with the nearly unlimited risk of violent escalation' (78) [3].

The first section of our article develops the effects-based logic for cyber war. Although US doctrine incorporates this approach, application of basic deterrence theory enables us to develop a more nuanced effects-based doctrine. We distinguish between counterforce and countervalue cyberattacks, and explore the implications for retaliation. Furthermore, we argue that a potentially important difference between kinetic and cyberattacks should be included in a sophisticated analysis: even if cyber and kinetic attacks are expected to inflict the same damage, there may be much greater uncertainty about the effects of the cyberattack. This difference in uncertainty/predictability has a variety of implications for cyber doctrine.

The second section addresses the possibility that saliencies exist, or could be established, in the cyber environment that would require the USA to modify the basic effects-based approach. For example, a retaliatory attack that inflicts extensive economic damage but no physical damage is likely to be understood differently from an equally costly attack that does inflict physical damage; there is likely a salient difference between economic and physical damage. Consequently, the USA should not envision these attacks simply as equally damaging. We conclude that a cyber doctrine that fails to incorporate saliencies risks overlooking the escalatory potential of certain retaliatory attacks. More specifically, there are likely to be situations in which kinetic retaliation will be more escalatory than a comparably costly cyber response. The third section explores possible norms that could constrain US doctrine. We suggest that development of a norm prohibiting cyberattacks against critical infrastructure, including even limited cyberattacks, is likely worth pursuing. We then discuss the possibility of an arms control arrangement in which major nuclear powers agree not to plan or launch cyberattacks against each other's nuclear command and control.

## Logic and implications of the basic effects-based approach (This section draws on [4])

### Deterrence basics

As others have argued, cyberattacks can be understood in the context of rivalry between states [5]. States may use cyberattacks to disadvantage their rivals; those rivals may sometimes retaliate. States may also seek to balance against other states by using cyber offensive operations where they feel that this provides them with a specific advantage. All this suggests that the logic of deterrence may apply, at least in principle (although as many have argued, questions of attribution will at the least complicate many deterrence based arguments). The key question then for US policy revolve around which deterrent threats or actions will be effective in preventing other states from behaving in ways that the US does not want, without leading to forms of escalation that are not in the US interest. Answering this question means, in the first instance, understanding what other states care about, and which threats are hence most likely to deter them without leading to unwanted escalation.

Our discussion here assumes that the state can determine who launched the attack; that is, we are putting aside the standard attribution problem. Given that our focus is on what type of retaliation the state should threaten and/or launch, little is lost by assuming that the state knows the identity of the attacker. In contrast, whether the target state would be able to determine that the state, not some other actor, retaliated could be an important factor influencing the state's choice between cyber and kinetic retaliation. Although we touch on this issue at the end of the section, it deserves more attention.

An effects-based approach starts from the claim that states care primarily about the 'extent' of damage that is inflicted, and not about the 'means' by which this damage was inflicted. For example, if an adversary undermines the functioning of a dam and causes severe flooding, it matters little whether the adversary employed a cyber weapon or a kinetic weapon—the state suffers the same damage and, more generally, the same costs.

If this perspective is correct, the state should envision its adversary in the same way. Deterrence then depends upon the ability to inflict damage against the adversary and/or to deny the success of the adversary's attack. As such, the adversary should care little whether these effects are achieved via cyber or kinetic means. Thus, to first order, whether the adversary is deterred should depend on its anticipation of effects/damage, not on the means via which the state promises to achieve them (This deterrence logic has a direct parallel in the applicability of international law to cyberattacks; see [6].).

The preceding points have direct implications for both deterrence by punishment—threatening costs—and deterrence by denial—threatening to defeat the adversary's attack [7]. 'Deterrence by punishment' relies on 'countervalue' attacks—that is, attacks against targets of inherent value as opposed to military targets, which are valued because of their ability to perform military missions. Valuable targets include a state's people, possibly its leadership, its economy, and related, the infrastructure that supports the state's people and its economy. Deterrence by punishment will succeed if the adversary believes the threatened costs are sufficiently large and sufficiently likely to be inflicted. In the nuclear realm, holding the adversary's cities hostage—that is, vulnerable to retaliation—is considered the basic requirement for deterring the adversary's nuclear attacks against one's own cities. The strict parallel in the cyber realm would be to threaten cyber retaliation that would

inflict comparable damage against the same type of targets that the adversary had attacked with cyber weapons.

The effects-based logic suggests, however, that we need to scrutinize this parallel much more carefully. From this perspective, there is no obvious reason that the USA 'needs' to deter countervalue cyberattacks with the threat of cyber retaliation. Because deterrence works by threatening costs with sufficient credibility, not by threatening specific types of attacks, this type of retaliation-in-kind is not necessary for deterrence to be effective.

There are many examples from the Cold War that are consistent with this basic point. For example, the USA relied on tactical nuclear weapons to contribute to its ability to deter Soviet conventional land attack. These weapons would have inflicted more damage than conventional weapons, but the point here is that the USA did not rely solely on retaliation-in-kind. The USA has retained the option of employing nuclear weapons to deter biological weapons attacks, among other reasons because the USA does not possess biological weapons. Studies of other US options for deterring biological weapons attacks have identified a range of conventional options, including invading the attacker's country [8]. All of these arguments are grounded in an effects-based logic. There does not appear to be a first-order reason that the USA should not rely on the same logic in planning to deter countervalue cyberattacks (However, as explore in the following section, there are considerations that are consistent with an effects-based approach that should constrain US retaliatory options to certain types of cyberattacks.).

If the USA wanted to make clearer that it was threatening/attempting to inflict comparable damage (for example, to avoid further escalation) via kinetic retaliation, it could attack targets that were similar to those its adversary had destroyed with cyberattacks. For example, if the adversary's cyberattack had destroyed part of the US electric grid, oil refineries, and or pipelines, the USA could retaliate against these infrastructure targets in the attacker's homeland. Alternatively, the USA could choose to threaten a type of damage that was quite different from that inflicted by the cyberattack. For example, except when facing a major power, the USA could threaten to invade the attacker's country or impose a new regime, if the country launched an extremely destructive countervalue cyberattack against the USA. These costs would be very different from those imposed by the adversary's cyberattack, but the costs do not have to be of similar types for an adversary to be deterred. In terms of the basic effects-based approach, the key consideration for the United States should not be whether to respond in kind—either in terms of means or targets—but rather which threatened response is likely to be most effective. As we explore below, this will depend on a variety of considerations, including credibility, expected effects, predictability of effects, and the availability and vulnerability of targets.

Whether the US retaliation should be proportional is an important question in deterrence theory, but not a central issue for the choice between kinetic and cyber retaliation. Threatening to inflict much greater damage than one's state suffered can be an effective deterrent, if the threat is highly credible. However, in some situations, threats that are 'too' large will lack credibility; threats to inflict a smaller amount of damage could therefore provide the more effective deterrent. From the basic effects-based perspective, whatever amount of damage was best for deterrence could be threatened by either cyber or kinetic means.

For all except possibly the most devastating cyberattacks, the USA would be able to inflict comparable damage with kinetic attacks against critical targets in the adversary's homeland. Depending on the nature of the adversary's economy and the extent to which it depends on vulnerable information networks, the USA might also have the ability to inflict comparable damage with a cyberattack (An important issue that we turn to later is the relatively uncertainty of the damage that would be generated by the two different types of attacks).

There is substantial disagreement on the damage that a sophisticated adversary could inflict with a cyberattack against the US homeland. Some authorities argue that extremely high levels of damage are possible: 'While the immediate effects of cyberattack are unlikely to be comparable to the effects of weapons of mass destruction (for example, nuclear, chemical, or biological weapons), a large-scale cyberattack could massively affect the functioning of a society and lead to many indirect casualties. Conversely, it is possible to imagine that certain cyberattacks might be executed on a smaller scale and with a lower degree of lethality than might be expected if kinetic weapons were used for equivalent military purposes. Thus the policy implications of cyberattack have certain commonalities across the range from non-lethal engagements to wars involving the use of weapons of mass destruction.' (26) [9] Others argue that the threat of wide-scale cyberattack has been over-rated; holding for example that 'Cybersecurity is an important policy issue, but the alarmist rhetoric coming out of Washington that focuses on worst-case scenarios is unhelpful and dangerous. Aspects of current cyber policy discourse parallel the run-up to the Iraq War and pose the same dangers. Pre-war threat inflation and conflation of threats led us into war on shaky evidence. By focusing on doomsday scenarios and conflating cyber threats, government officials threaten to legislate, regulate, or spend in the name of cybersecurity based largely on fear, misplaced rhetoric, conflated threats, and credulous reporting (83–84) [10].' (See also [11]). We do not seek to adjudicate this argument in this article. In the absence of any evidence of a wide-scale cyberattack or attempted cyberattack having occurred, it is hard to be sure how wide the consequences would be, since large-scale damage would likely result (if it happened) from cascading effects, unknown interdependencies and other phenomena that are complex in the technical sense of that term. However, for the sake of analysis, we look at the 'possibility' that such an attack might occur.

A key potential shortcoming of kinetic retaliation must therefore lie in the adversary's assessment of US credibility—that is, the adversary's assessment of the US' capability and willingness to inflict retaliatory damage via kinetic attack (A second potential shortcoming, which is the focus of the following section, is that kinetic retaliation might cross an important saliency and therefore result in larger escalation). These shortcomings need to be compared to the credibility challenges inherent in cyber retaliation, which are substantial.

To lay the groundwork for this comparison, we first consider potential barriers to making cyber retaliatory threats credible. Cyber retaliation may in general be less credible than kinetic retaliation, because a state will have greater difficulty demonstrating its cyberattack capabilities. States can demonstrate their conventional and nuclear capabilities by developing, testing and deploying forces, demonstrating their effectiveness against relevant types of targets, and engaging in training and exercises, all of which are observable (to varying degrees) by its adversaries. In contrast, the adversary will have far less evidence of the extent and effectiveness of US offensive cyber capabilities. Not only are they entirely invisible, but they may be untested against adversary systems, leaving the adversary with some doubt about the effectiveness of US capabilities, and in turn about the credibility of its threats (Under one interpretation such problems mean that the Snowden revelations may actually have benefited the credibility of US cyber threats, by providing

information about the extent of US offensive cyber capabilities; see [12]). Testing cyber weapons against the adversary's systems, especially ones that it views as especially valuable and important, would be risky because if detected the adversary would likely view the test as highly provocative. In addition, testing a cyber weapon could reduce its future effectiveness, because an adversary that detects the test will also be alerted to the vulnerability that the attacker is planning to exploit. Doubts about the attacker's offensive cyber capabilities could be further increased by the potential limitations of relying on one-shot and/or bespoke weapons such as zero day exploits (57–59) [13] (It is true, however, that the US reputation for being highly capable in IT, and in cyber more specifically, could help to reduce this potential credibility problem). Thus, conventional responses will often have an advantage in terms of the adversary's ability to assess US capabilities to inflict costs.

How does the credibility of kinetic retaliation compare? First, the adversary might doubt the appropriateness of a conventional response, believing that retaliation-in-kind is the most obvious response. Although this is a natural consideration, an effects-based perspective suggests that it should be largely discounted: why should the means that the US employed to inflict damage influence the adversary's assessment of US credibility for inflicting a given level of damage? If anything, the analysis so far suggests that conventional retaliation has important advantages. For the reasons we sketched above, the adversary will have less doubt about the US ability to launch an effective conventional attack than an effective cyberattack. In addition, uncertainty about the scope of the effects that a cyberattack would inflict—especially the possibility that it would do far more damage than intended—could make the US leaders reluctant to order such an attack. Recognition of this complexity-induced reluctance could, in turn, reduce the credibility of the US countervalue cyber retaliation.

Yet a third factor that could favor conventional retaliation is the relative vulnerability of the USA to cyberattacks compared to conventional attacks. The USA is a densely networked society, with a rich variety of targets for countervalue cyberattacks. Some potential adversaries may not be so rich in cyber vulnerabilities. In this type of case, if the adversary believes that the USA expects retaliation-in-kind (which we have argued is not a clearly logical position), then the adversary would find the US conventional threats more credible, because the USA was less vulnerable to conventional retaliation than to cyber retaliation.

Second, the adversary might question whether the USA would be willing to escalate to conventional retaliation, if it believed that the USA believed conventional retaliation would lead to escalation to still more damaging attacks. Once again, from an effects-based perspective there is not an obvious reason for this belief. It is possible that more subtle understandings of escalation thresholds or steps in an escalation ladder could support this concern. In the following section on saliences, we explore whether this type of distinction might exist between the specific effects of cyber and kinetic attacks, and whether the USA has the ability to influence these understandings.

To sum up, the effectiveness of the US deterrent will be enhanced by leveraging both its (known) kinetic prowess and its (partly unobservable) cyber prowess to make deterrent threats. Precisely how to draw upon both sets of assets is complicated: not specifying in advance which it might use increases the range of retaliatory options the adversary must take fully into consideration; on the other hand, making specific threats if specific types and/or levels of fighting occur puts the USA' reputation on the line, which can contribute to the credibility of specific threats. One thing that is clear, though, is that the United States should rely, at least partly, on its kinetic options, as it already does.

'Deterrence by denial' works by an entirely different logic: in this approach, the USA deploys capabilities to convince its adversary that the probability that its attack will succeed is low; this reduces the adversary's expected benefits of the attack and can therefore result in successful deterrence. Even more than deterrence by punishment, the type of scenario plays a critical role in evaluating the choice between cyber and conventional denial.

If considering a cyberattack that does not inflict physical damage, then the denial capability will typically be cyber; that is, because the attack is against cyber systems, the way to defeat it will ordinarily be some type of cyber capability, whether defense, redundancy or an offensive cyberattack that disrupts the adversary's attack (although some forms of physical protection, such as, most obviously, air gaps, may also be efficacious).

On the other hand, if considering a cyberattack against US military capabilities, US options are then quite different. Deterring cyberattacks in isolation is probably not the key to deterring this type of attack. Both the USA and its adversary are likely to envision counter-military cyberattacks as an integral part of their overall conventional fighting capability. Within types of weaponry and warfare, the USA has traditionally distinguished between conventional and nuclear warfare, and also made distinctions concerning chemical and biological weapons. In contrast, in the context of counter-military attacks, cyberattacks should not be considered a different type of warfare. Instead, counter-military cyberattacks should be viewed as a component of conventional warfare.

This would be in line with current categorizations, which for example include electronic warfare assets as an element of conventional capabilities. Similarly, imagine a cyberattack that damaged US command and control capabilities. Why should the USA response to this attack, or its deterrent threat that is designed to prevent the attack, be different if the damage is done by a kinetic attack than by a cyberattack?

If the preceding line of argument is correct, then the challenge the USA faces in deterring counter-military cyberattacks is to be able to deter the adversary's overall conventional attack, including the offensive cyber capabilities that would be a component of this attack. This overall deterrent will depend on relative US cyber capabilities, including both its ability to defend against the adversary's cyberattacks and its ability to use offensive cyberattacks to weaken its adversary's overall conventional capability. But, deterrence will depend still more broadly on how US conventional capabilities compare to its adversary's. The adversary could be deterred from launching a conventional attack, including its counter-military cyber component, if the USA has the ability to win a conventional conflict, even if its adversary enjoys a cyber-advantage. And, more in line with standard worries, an adversary that enjoys a net advantage in counter-military cyber capabilities might not be deterred, even if US conventional forces are otherwise clearly superior. In any event, the basic point here is that the impact of cyber capabilities on deterrence has to be understood in terms of their net impact on US overall conventional capabilities.

Given that expectations about the combined overall impact of conventional and cyber capabilities will determine the effectiveness of the US ability to deter by denial a conventional war, including cyberattacks, the USA should choose the mix of conventional and cyber capabilities that will have the best prospect for defeating the adversary's attack and, closely related, for deterring that attack in the first place. The proper mix of US conventional and cyber assets is likely to vary across specific conventional war scenarios.

Once again, relying heavily on conventional capabilities has one clear advantage—the USA is likely to have greater confidence in these capabilities. As a result, its adversary may view conventional forces as a more convincing deterrent.

## Deterrence implications of uncertainty about the effects of cyber retaliation

One important difference between conventional and cyberattacks, which we have only touched on so far, deserves more attention—the uncertain effects of cyberattacks. There is a general belief in the literature that cyberattacks have more unpredictable consequences than conventional attacks [9, 13]. This unpredictability reflects the nature of cyberattacks, which typically involve attacks on complex architectures of software, which are connected to other computers via the Internet. This unpredictability reflects three aspects of a cyberattack and cyberspace more broadly.

First, the complexity of the target software itself could render an attack unpredictable simply by obscuring what would happen when the software systems is interfered with or disrupted. Second, because most computer systems are connected to other computer systems via the Internet, some kinds of attack could spread across these computers. The complexity of each system and how they are connected mean that it is hard to make predictions about the extent and speed of spread and the impact on each computer. Third, corruption of computers could generate physical effects that cascade well beyond cyberspace and are themselves difficult to predict. For example, a cyberattack against computers that control a limited portion of the electric grid could lead to much more far reaching damage, if local outages themselves create other outages across the grid in a cascading process.

A well-known example of unpredicted spread of a cyber virus is Stuxnet: the attack ended up infecting many 'innocent' computer systems in Iran and elsewhere, although it did not inflict physical damage beyond the Iranian nuclear complex [14]. Unconfirmed reports suggest that other cyberattacks have had unexpectedly extreme consequences (such as briefly taking out an entire country's Internet access for a period, more or less by accident) [15]. Reports suggest that the unpredictable collateral damage of a large scale US cyberattack on Iran played an important role in war planning discussions [16]. It is also plausible that attacks that were intended to have large-scale consequences have fizzled or failed, because the targeted system did not respond in the predicted ways. Such failed attacks will often be invisible to everyone except the attackers.

Reflecting this overall uncertainty, the variance in the damage inflicted by a cyberattack is likely to be greater than for a kinetic attack (This will, however, vary somewhat with the type of kinetic attack. For example, a precise kinetic attack against a portion of a state's electric grid could generate a cascade of blackouts comparable to a cyberattack that damaged the same portion of the grid). In other words, the distribution of damage that would be inflicted by many types of cyberattacks is likely less tightly clustered around the hoped for/planned damage than for many types of kinetic attacks.

Uncertainty about the damage a cyberattack would inflict could make kinetic threats more effective deterrents than cyber threats. The effectiveness of deterrent threats depends on a state's ability to carry out the threat: deterrence by denial is less likely to succeed if one's adversary believes a threatened response is unlikely to achieve its military objective; and deterrence by punishment is likely to fail if the adversary doubts the state's attack will inflict the promised damage.

Moreover, except in an all-out war, a state would want to be confident that its attack would not inflict more damage than intended, because doing so could increase the probability that the adversary would escalate still further. For both of these reasons—credibility and escalation control—cyberattacks appear to be less effective deterrent tools than are kinetic attacks.

However, there may be circumstances under which the unpredictability of a cyberattack could make it *more* attractive than a kinetic attack. Building on Schelling's discussion of 'the threat that leaves something to chance', (chap. 8) [17] (171) [18], we might imagine that an 'attack that leaves something to chance' could be an effective deterrent under some circumstances. The threat that leaves something to chance promises some probability of a very costly outcome, in a situation in which the decision to carry out the action is not under the control of the threatener. In situations in which the threatener would also be hurt by the action that inflicts very high levels of damage, or by the adversary's likely response, the threat that leaves something to chance can be more credible, because the threatener may be willing to run some probability of suffering the damage, but not unwilling to suffer it with certainty. It would also be more credible when the threatener was unable to turn off the threat; otherwise the target of the threat might wonder whether the threatener might pull back its threat.

The 'attack that leaves something to chance' would work by a related, but somewhat different logic. An attacker that launched a cyberattack that might impose extremely high costs would demonstrate its resolve (that is, the extent of her interest) in prevailing in the conflict and thereby gain a bargaining advantage in a limited war. Although the attacker would be unwilling to inflict the costs with certainty, due to high probability of costly escalation, she is willing to take a chance/run the risk of inflicting these costs. And, by the nature of the effects of the cyberattack, once the attack is launched, the attacker cannot prevent the worst outcome from occurring, which should reinforce the target's judgment about the attacker's resolve. As a result, in certain situations, a cyberattack that is on average expected to inflict the same amount of damage as a kinetic attack could be the more effective tool for compelling, intrawar deterring and bargaining.

Interestingly, the deterrent value of threatening an attack that leaves something to chance (to be distinguished from actually launching such an attack) is not so clearly greater than the comparable kinetic threat: although the target would recognize the possibility of suffering greater than the average damage, she would also be aware of the possibility that the cyberattack might inflict less than the average damage. Risk-averse states would see the cyber threat as more costly, while risk-acceptant states would see it as less threatening than the comparable, more certain, kinetic threat.

Finally, it is possible that the ambiguities and uncertainties associated with cyberattacks may sometimes have another advantage. Policy makers and analysts have devoted enormous attention to the 'attribution problem'—the difficulty of attributing cyberattacks to their attackers. This could be a major problem for deterrence, yet it can, under some circumstances, be a blessing. It provides states with greater freedom of action in choosing how or whether to respond to an attack. Consider the difference between a physical attack and a cyberattack that destroys or degrades an important asset belonging to an adversary. It will be difficult for the adversary to ignore a physical attack without appearing feckless or weak. If it is capable of responding, it will likely have strong incentives to do so, in order both to demonstrate resolve to possible attackers and to avoid criticism from domestic audiences. In contrast, the state will have more leeway in deciding how to respond to certain cyberattacks. Even if

the state knows who the attacker is, the attacker does not necessarily know that it knows, nor do other states necessarily know that it knows. Thus the lack of common knowledge and difficulty of attribution reduces the state's ability to deter cyberattacks, but for closely related reasons could allow the state to avoid a potentially costly response that it would prefer to forgo. The state can act as though it does not know whom the attacker was and hence decline to retaliate without doing great damage to its reputation for resolve.

## Salience and focal points

The basic effects-based approach provides a relatively simple framework for thinking carefully about offensive capacities and deterrence in cyberspace. The approach, however, does not provide a sufficiently rich description of how states are likely to actually understand cyberattacks, especially compared to other types of attacks. It could be that states do not view all equally damaging attacks as equal. To understand this possibility, we turn to the concepts of focal points and saliences, which capture the implications of states' shared understandings of actions, and in turn their reactions to others' actions and their expectations about how other states will react to their actions.

Imagine a scenario in which the New York Stock Exchange suffers a cyberattack that prevents stock trading for a period of weeks, thereby inflicting quite significant damage on the US economy. Imagine further that the USA responded with a kinetic attack aimed at the central business district of the adversary's capital, which inflicted damage to property equivalent in value to the economic damage of the cyberattack, without any people being wounded or killed. Under this scenario, would others (adversary and other states) consider the US retaliation to be commensurate, or alternatively escalatory?

An effects-based account would clearly predict that this response would be considered to be commensurate. Our immediate intuition, however, is that this prediction is wrong. Other states, including the adversary, would likely consider the retaliation to be a substantial escalation.

Yet there are other circumstances in which our intuition suggests that responding with a kinetic attack would likely not be viewed as escalatory. Imagine for example, that the initial attack was on military rather than civilian assets—say, for example, an adversary attacked and disabled an important military system using cyber means (as Israeli forces reportedly did as part of an air raid on Syria, disabling radar) (1–9) [19]. Here, we suspect that a kinetic response aimed at a military system of similar importance would not be regarded as escalatory, or at the least would be seen as ambiguous.

The analytic challenge posed by intuitions like these is knowing whether they are widely shared. Intuitions may differ importantly from person to person and from state to state, which increases the probability of misunderstanding and hence of unintended escalation. One valuable intellectual tool for exploring, and possibly clarifying these intuitions, is Thomas Schelling's arguments about salience and focal points. These can help both to sharpen analysis among observers and increase predictability in inter-state relations. During the Cold War, analysts believed that developing a common vocabulary and understanding of how actions might be interpreted could contribute to stability during crises and limited wars.

Salience results from focal points, which serve as an implicit solution to coordination games. In the context of a specific situation, focal points tend to possess some kind of 'prominence, uniqueness, simplicity, precedent, or some rationale that makes them qualitatively differentiable from the continuum of possible alternatives'(chap. 30, quote at 70). [17]. If one imagines a coordination problem in which actors need to converge upon one of many possible solutions in order to coordinate properly, then actors will plausibly turn to commonly shared focal points in order to predict how other actors might behave, and hence reach an equilibrium. In Schelling's famous example, students who had to decide where to meet in New York at a given time, without any information as to which of the many thousands of possible locations in New York they should meet at, are likely to converge on Grand Central Station as a plausible location. Here, they draw on forms of information that are external to the intrinsic strategic situation in order to successfully resolve it. An actor is drawn to the focal point solution because of her belief that (i) other actors are likely to view the feature as a focal point, and (ii) other actors are likely to expect that the given actor also sees the features as a focal point, and so on, which creates common knowledge and hence generates converging expectations.

Such information might come from prominent features of the landscape such as rivers, the political status quo, and differences between types of weapons—for example, conventional versus chemical versus nuclear [17], labels associated with specific strategies, culture and institutions (3–30) [20, 21] or other such distinguishing aspects of the situation that actors face which are not givens of the strategic structure of the situation itself (Another type—skewed distributions in which some solutions are mentioned far more often than others is explored in [22]). These are the various factors that make focal points *salient*, enabling actors to converge on common understandings in ambiguous situations.

Notably, salience and focal points can play an especially important role in deterrence, compellence, escalation control and limited war. Saliences provide distinctions between specific categories of action, some of which are viewed as escalatory while others are viewed as restrained, some of which are viewed as requiring a harsh response while others are viewed as tolerable, some of which are expected while others are surprising. Possibly the clearest saliency in current weaponry and war is between conventional weapons and so-called weapons of mass destruction, with the conventional-nuclear divide being the sharpest. As both rationalists and constructivists have observed, states draw a sharp distinction between conventional and nuclear weapons. The distinction is not based entirely on the damage that a nuclear weapon would inflict. The USA can build nuclear weapons that would do less damage than would the largest conventional explosives and far less damage than the overall damage inflicted by large-scale conventional bombing.

Instead, as Schelling described it, the understanding that nuclear weapons are 'simply different and generically different' is based on an argument that 'emphasized bright lines, slippery slopes, well defined boundaries, and the stuff of which traditions and implicit conventions are made' [23]. Crossing the nuclear saliency by using even a single small nuclear weapon is believed to greatly increase the probability of further nuclear use and possibly of massive nuclear war. Among other reasons, this is because once nuclear weapons are used there may not be any 'natural' place for the warring parties to stop.

If settled and relevant cyber focal points exist, then states will look to these saliencies to predict how other states will interpret a cyberattack. This may render some cyberattacks more escalatory and/or threatening than comparably damaging kinetic attacks, and vice versa. Hence, if relevant saliences exist in cybersecurity, then an effects-based approach to deterrence is incomplete: actors may respond to an attack in ways that the simple effects-based account

would not predict, because the attack crosses a salient dividing line. This, for example, could explain why actors might converge on agreement that a kinetic response to a cyberattack on civilian infrastructure, even if no one were killed, would be escalatory. It might also explain why actors could agree that a kinetic response to an attack on military infrastructure would not be escalatory. Clearly, working out the contours of the perceived landscape of possibilities that make one distinction focal and the other not, is essential for designing US cyber strategy.

Consider a few possible reasons why a kinetic response to a cyberattack might be considered escalatory. First, it could result from the belief that cyberattacks and kinetic attacks are fundamentally different in *kind*, such that one is considered fundamentally acceptable, and the other is considered non-acceptable. If this were generally accepted, then the effects-based doctrine that we outlined initially would be more or less useless, since it would be undermined by an understanding that there is a crucial qualitative difference between all cyber and kinetic attacks. However, if we are not alone in our intuition that a kinetic response to a cyberattack on military infrastructure is non-escalatory, it would seem unlikely that such a sweeping and general distinction applies.

Second, it could be that the focal point turns on a perceived difference between physical and non-physical *damage*. This would imply that forms of cyber attack might be considered equivalent to kinetic attacks when they do direct physical damage. For example, an attack on a dam's control system that created major flooding might be seen as equivalent to a directly kinetic attack that produce the same flooding. If this were the key distinction, then we might expect differentiation between different kinds of cyber attacks, depending on the type of damage that they inflicted. For example, attacks that damaged information systems or electronic commerce would be viewed as non-comparable to kinetic attacks, while cyber attacks that did direct physical damage would be viewed as comparable.

Third, the distinction could turn on whether the cyberattack inflicts easily observable damage. The losses from a crippled stock exchange are plausibly less visible than the losses from a kinetic attack that does immediately observable damage to buildings and infrastructure. Here, cyberweapons that did observable damage (which might be physical, but might also involve purely virtual effects that had easily observable consequences) might be viewed as equivalent to kinetic attacks that were equally visible.

Fourth, we expect there will likely be a relevant distinction between a cyberattack that inflicts physical damage on military assets and one that inflicts physical damage on civilian assets. As we have noted, we suspect that kinetic counter-military responses to substantial cyberattacks on military targets are less likely to be viewed as escalatory.

Fifth, a likely distinction is between cyberattacks that kill people and those that do not. Cyberattacks would not kill people directly, but could result in physical damage that would then kill people. Even visible physical damage that is very costly might be considered less escalatory than an attack that kills people but is otherwise not very costly. It is less clear whether cyberattacks that impose material costs on people—for example, depriving them of electricity—but do not kill anyone would be viewed as more escalatory that attacks that inflict great financial harm but do not have immediate material consequences for people's lives. This set of distinctions applies as directly to the differential effects of kinetic attacks; that is, the distinction is not special to cyber.

Still other saliencies may exist. There is possibly a distinction between attacks that occur during wartime and peacetime. Kinetic attacks on civilian infrastructure may be less likely to be viewed as escalatory if states are already involved in armed hostilities. There is also the possibility that states will view cyberattacks that temporarily interrupt the operation of systems—for example, an attack that takes down the electric grid but does not permanently damage it—as less escalatory that a kinetic attack that does permanent damage, even if the two attacks inflict equal overall economic costs.

Given that states' understandings of cyber warfare are at an early stage, the USA should consider whether there are possibly feasible focal points that it would like to help establish. Because we expect that agreed upon saliencies have the potential to reduce undesired escalation in wars that involve cyberattacks, establishing shared understandings could be in all states' interests. The USA should also consider whether other actors—whether adversaries, allies, or non-state actors—may also be seeking to establish focal points, and what those focal points might be. Not all focal points will be desirable. Some may limit US freedom of action by making certain types of attacks more escalatory than they would be if the focal point did not exist or, closely related, if the USA was known to reject the contested focal point.

This discussion raises the question of how, if at all, the USA can contribute to the establishment of focal points in cyber war. One approach may be negotiations or possibly official dialogues in which states share which saliencies, if any, they believe operate or can operate in cyber war. But active efforts to build focal points need not be limited to negotiations and discussions of the issue. In fact, it may well be that threatened actions, actual actions (and non-actions) and the interpretations of actions will contribute more to the establishment of saliencies.

One potential source of influence on focal points could be US cyber doctrine. Current US doctrine makes clear that the USA retains the option to employ a kinetic response to a cyberattack. In effect, the doctrine denies that, at least in broad terms, there is a salient difference between cyber and kinetic attacks. This preservation of flexibility over responses to cyberattacks has not received harsh criticism from other states, or even any sustained opposition from civil society. It is likely too early in the cyber age to know whether this reflects acceptance and recognition of the lack of a broad salient distinction between cyber and kinetic attacks. A broader evaluation and discussion of salience in cyber war—within the US government, with experts outside the government, and between relevant governments—might help establish greater clarity before the test of war brings its own form of clarity to these issues.

In evaluating its possible interest in the creation of saliencies, the USA should consider the disadvantages of its current cyber doctrine. By reserving for itself the right to retaliate against cyberattacks using non-cyber means, the doctrine provides other states with some justification for behaving in the same way, employing conventional means to respond to US cyberattacks. For example, if Iran had been capable of launching a kinetic attack against the US homeland in retaliation for the physical damage that the Stuxnet virus inflicted on its nuclear complex, would the USA be willing to accept that this was a reasonable form of retaliation or would it have understood it as highly escalatory? (We don't mean to suggest by using this example that Iran did not employ kinetic retaliation against the USA or its allies due to this distinction. Among other possibilities, Iran might have been deterred by the possibility of US escalation). At the least, it is more difficult for the USA to complain about other states responding to cyberattacks with kinetic force if it reserves the same option for itself.

A second potential source of influence on understanding of focal points will likely be US (and other states') actions in response to

large-scale cyber attacks that variously inflict economic, physical or military damage. To the best of our knowledge, there has not been a kinetic response by a state to a cyber attack. This, however, provides relatively little information about existing saliencies because the USA has not suffered a cyberattack that was sufficiently large and costly to initiate what would traditionally be considered an interstate war, in which kinetic retaliation might appear to be the 'natural' response. The USA has suffered cyberattacks below this level and employed non-kinetic forms of retaliation. For example, it has indicted Chinese 'military hackers' for hacking, espionage and other offenses [24]. It also imposed sanctions against North Korea after Sony's servers were hacked, and according to one prominent member of Congress, cut off North Korean access to the Internet for a period of time [25]. Drawing insights from these cases about the existence of saliences is further complicated by the possibility that a saliency was crossed, but the USA was deterred from inflicting the more costly or escalatory retaliation that might then have been appropriate (On the USA being deterred, and also for criticism of its mild responses, see [26]).Whether a saliency is crossed is only one factor in a state's decision to escalate. Escalation might not be the state's best option, if the risks are too high. Thus, a state's reactions may not map neatly into the crossing of saliencies.

The preceding discussion leads us to the following conclusion: US doctrine may have to take saliencies and focal points into account. The basic effects-based argument implicitly assumes that all effects can be aggregated into a single value. Essentially, each type of damage, including human lives, can be given a dollar value and all of the costs can be added together to determine an attack's total cost/effect. Appreciating the potential impact of saliencies requires us to reject this approach. Instead of aggregating across types of damage, we may need to identify different dimensions along which states and individuals distinguish types of damage and then be cautious about ranking the overall severity and information content of an attack. Different types of damage may simply be different—physical or not, human lives lost or not, easily observable or not, military or not, temporary or permanent—all of these may influence how an adversary understands an attack. When they do, an attacker will need to incorporate these dimensions into its decision about what type of threats to make and how to retaliate, if deterrence fails. As a result, in some but not all situations kinetic responses to costly cyberattacks will be inappropriate, or at least more escalatory than the basic effects-based approach would indicate.

## Norms

Another potential limit to the effects-based doctrine involves norms. Norms are internalized, and at least to some degree do not involve a means-end distinction, which makes it nearly impossible to incorporate them into an effects-based argument. Focal points and salience operate through a strategic logic that can be carried through by rational ends-focused actors. In contrast, internalized norms are 'nonconsequentialist'—that is to say, that they involve judgments as to whether actions are innately appropriate or inappropriate, regardless of their consequences [27].

For example, the previously discussed distinction between nuclear and conventional weapons is plausibly not only a focal point, but also a partly internalized norm. The first use of nuclear weapons is regarded as a 'taboo' that can be violated only under extreme circumstances. The developing norm was recognized early in the nuclear age and has become more deeply established with time [28]. The animus against nuclear weapons stems not only from logic but also from 'moral discourse about nuclear weapons' that was often viscerally hostile to the effects-based logic of deterrence (372) [28]. Nuclear weapons came to be seen as not only profoundly different from ordinary weapons, but in addition their first use can be viewed as unacceptable.

The question we need to address is whether there are norms of cyber war that could or should place limits on US cyber doctrine. Put another way, are there offensive cyberattacks that the effects-based approach supplemented by saliencies would prescribe, or at least not proscribe, that the USA would be unwilling to launch because they are normatively inappropriate? The answer—at least for the moment—appears to be 'no'. As discussed above, there may be cyber saliencies that the USA should not cross because doing so would unduly increase the probability of escalation. In contrast, there do not appear to be offensive cyberattacks that the USA believes it would simply be 'wrong' to launch, except for those that violate standard laws of war.

The USA has engaged in some informal norm-building efforts in cybersecurity. In part this is because one of the key alternatives for limiting an adversary's capabilities—formal cyber treaties—would usually be exceedingly difficult, likely impossible, to verify. Unlike nuclear and conventional weapons, which states can often effectively monitor, cyber capabilities are largely invisible to the outside observer. In addition, many of the capabilities an adversary could use to launch an offensive cyberattack could also be used to defend against one. Once an attack was launched, the state might not be able to identify the perpetrator with a high confidence, and even if it could, might not be capable of proving it to other states. Finally, even if all of these barriers could be overcome, it is far from clear that the USA would be willing to trade away its offensive cyber capabilities in return for its adversaries foregoing theirs. For all of these reasons, the USA has not focused much energy on achieving formal treaties (which could have normative consequences as well as legal consequences and associated sanctions), but has instead looked to informal and quasi-formal understandings that do not rely on formal compliance mechanisms, but instead on the identification of appropriate standards, and the shaming of those who do not live up to these standards.

The most visible exercise in attempted norm building is not in cyber-offensive operations as such, but in cyberexploitation—cyber operations that are aimed at extracting information rather than paralyzing, degrading or damaging assets [9]. The USA holds that there is a basic distinction between purely commercial cyberexploitation—securing commercial secrets that are then shared with favored domestic businesses—which it considers illegitimate, and regular cyberexploitation—gathering of information relevant to national security—which the USA considers legitimate. Other countries, including prominently China, have disagreed. This disagreement may partly reflect different relationships between the state and the private sector: the USA does not have a history of strong direct state involvement in directing commercial activity; in contrast, many other countries do not have such an arms-length relationship between the state and the private sector, which would help explain why their perspectives differ.

China is not the only important example. France, too, has acquired a reputation for flexibility in the sharing of commercially valuable information with businesses that sometimes were formerly state owned, and still retains very strong state connections.

The result has been serious disagreements between the USA and China. As already noted, the USA, lacking apt multilateral international instruments to express its displeasure, turned to domestic law enforcement, for example seeking indictments against Chinese

nationals that it claims have been involved in commercial spying and threatening sanctions. While these indictments are highly unlikely ever to result in successful prosecutions, they carry some weight in signaling US normative priorities and in shaming China.

This and other pressure likely led the USA and China to reach an understanding under which China has agreed not to 'conduct or knowingly support cyber-enabled theft of intellectual property, including trade secrets or other confidential business information, with the intent of providing competitive advantages to companies or commercial sectors' [29]. Although the informal agreement lacks explicit enforcement mechanisms, it has apparently led to some reduction in cyberexploitation against US companies.

The ability to reach an informal agreement likely resulted partly from the process it creates, of repeated discussion between the USA and China (and perhaps other actors) about what the norm actually entails. The informal agreement requires both countries to consult with each other regularly about enforcement activities, creating a 'high level joint dialogue mechanism' [30, 31].

This reduction may be the result of norms in action. According to this explanation, U.S. legal action shamed China and led to a shift in China's public position on commercial cyberexploitation [32, 33]. There is, however, an alternative explanation, which holds that China is employing external pressures created by the USA to gain domestic control of actors who are pursuing their own economic interests with inadequate regard for China's overall strategy [34, 35]. The available evidence is ambiguous, and could be interpreted as supporting either of these explanations (or perhaps some combination).

There is even less normative agreement regarding cyberattacks. United Nations reports have agreed that international law applies to cyberspace, but have provided little guidance on their implementation. The reports do not address the application of international humanitarian law to cyberspace [36].

To end our discussion of norms, we identify a norm that the USA should consider promoting. While we believe this norm would be potentially valuable, we are not advocating for it; instead we are encouraging further exploration.

The norm would be a prohibition on attacking critical infrastructure (For a similar and more developed recommendation (24–26), see [37]; see also [38] who focuses on the possibility of an arms control agreement, not a norm (81–86)). The rationale is that such an attack might inflict crippling economic damage that could far exceed feasible conventional attacks. The publicly available scholarship disagrees on whether large-scale counter-infrastructure cyberattacks could have such severe and even crippling consequences for civilian infrastructure [9, 10, 11]. Without taking a position on the destructive potential of such attacks, we can make a qualified argument. If such attacks are a plausible danger, then the USA should consider supporting development of a norm against cyberattacks that target infrastructure that is critical to the operation of states, including electric grids, oil refining facilities, and backbone financial networks. The USA and other countries appear to be moving in this direction. The United Nations' Group of Governmental Experts included among its 'recommendations for consideration by States for voluntary, non-binding norms, rules or principles of responsible behaviour' that 'A State should not conduct or knowingly support ICT activity contrary to its obligations under international law that intentionally damages critical infrastructure or otherwise impairs the use and operation of critical infrastructure to provide services to the public'. (8) [39]

A possible criticism is that a limited counter-infrastructure attack—for example, one that targeted only facilities in a geographically limited region—would not do truly catastrophic massive damage and therefore should not be subjected to this normative prohibition. However, a few rejoinders carry weight. To start, this type of limited attack could reduce the barriers to additional similar attacks, possibly leading to unlimited cyberwar. As with limited nuclear use, once the saliency is crossed there may not be a 'natural' place at which to reestablish tacit limits on countervalue cyberattacks. In practice, an unavoidable complication that is not present in the nuclear case is that the line between counter-infrastructure attacks and other cyberattacks may be less clear than the nuclear-conventional divide. Efforts to establish this norm would have to engage this complexity, among others. Another counterpoint is that a limited counter-infrastructure attack could result in more far-reaching damage than the attacker intended: an attack against a specific region could extend much further either via cascading damage that flows from the interconnectedness of critical systems or via the unintended spread of the cyber weapon itself. Thus, states should recognize that limited attacks against critical infrastructure are too risky and should stigmatize them. In the terminology of our earlier discussion, the 'attacks that leave something to chance' should be rejected as an unacceptable tactic in intra-war bargaining.

Another possible criticism is that states often violate norms and specifically that they have often violated the norms against harming non-combatants. Hence, one might argue that norms are effectively worthless. However, in wars of attrition, states have often not violated the norm against targeting civilians until late in the war, and violate the norm out of determination or even desperation to win [40]. If the same logic applied to cyber war, then a norm against counter-infrastructure attacks could contribute to delaying these attacks and possibly thereby avoiding them.

A norm against attacking critical infrastructure would augment the effects-based approach and would be grounded in the potential effects of such an attack. At least until recently, major powers were incapable of inflicting crippling damage against an adversary's critical infrastructure, and thereby its society, with conventional weapons before gaining control in a total war. At a minimum, this level of damage could not be inflicted quickly. By possibly making this option available, counter-critical infrastructure attacks appear to create a new danger, one that is sufficiently large that states should consider judging it an unacceptable action.

The prospects for developing this norm are improved by the typical interest of norm entrepreneurs in threats to people's lives, especially potentially large threats. The United States is hampered in many of its efforts at norm building by distrust and strong disagreement from the technology community and other states [41]. However, there is plausibly scope for agreement over norms against cyber-attacks that would result in significant loss of civilian life, as reflected by the agreement of states that the ordinary laws of war ought apply to cybersecurity. This could potentially be expanded into a set of norms against the use of cyber weapons against critical infrastructure, even though many of the deaths might result from indirect effects of the attack. Over time, if states observed this norm, it could become internalized, resulting in the delegitimization of cyberattacks against civilian infrastructure.

## Arms control

A rather different type of restriction on US cyber doctrine could result from an arms control agreement to forego certain types of

cyberattacks. We believe that the USA should seriously explore the possibility of an agreement that would prohibit cyber intrusions into the command and control (C2) systems of the major powers' nuclear forces (26–27) [37]. The effectiveness of a state's nuclear deterrent depends on its ability to credibly threaten retaliation, which requires not only that its force survive an attack, but also that its ability to launch those forces survives. Vulnerable C2 can undermine a state's nuclear deterrent and create dangerous dynamics during a crisis (On the vulnerability of nuclear command and control, and the dangers it can create, see [42]; for a recent analysis of these dangers in US nuclear strategy toward China, see [43].). A state, however, could believe that holding its adversary's C2 vulnerable could provide strategic advantages, especially if it was also able to target much or all of the adversary's nuclear force. Given the potential advantages of being able to attack the adversary's nuclear C2, but also the disadvantages, a state might be willing to forego the ability to launch this type of a cyber attack if and only if its adversary were willing to do so as well. Because intelligence gathering efforts would likely be indistinguishable from preparation for a cyberattack against nuclear C2, mutual restraint would almost certainly need to include both.

A state would likely only engage in this mutual restraint if it had a reasonable chance of verifying the adversary's restraint—that is, of detecting cyber intrusions into its nuclear C2 system and identifying the intruder. While the feasibility of detection and attribution is primarily a technical issue, one factor that would favor feasibility is timing: various types of preparation for a counter-C2 attack would almost certainly be required during peacetime; consequently, a country would likely have a substantial amount of time to inspect for intrusions. Finding a single serious intrusion would likely be sufficient to bring its own restraint to an end. The adversary's recognition of this likelihood could deter it from violating the mutual restraint on preparing cyberattacks against nuclear C2 (For an argument that even passive intrusion into command and control systems could be regarded as a grossly provocative action, see [44].).

An alternative to an arms control agreement would be a norm against nuclear C2 attacks. However, whether a norm against counter-C2 cyber would be valuable and can be developed is less clear. Regarding its value, one could argue that if verification is possible, then a norm is unnecessary; this has been the model for past arms control agreements. Regarding its feasibility, counter-C2 cyber capabilities would have to achieve a special status, one that makes them clearly more dangerous than other types of counter-nuclear and counter-C2 weapons. The USA has not forgone these capabilities and has built them into its war plans. Some features of cyber might distinguish it from these other weapons—most obviously, the greater uncertainty that cyber counter-C2 weapons might create about the vulnerability of an adversary's nuclear retaliatory capability. However, at most this is likely to be a difference of degree, not kind, which suggests the prospects for developing this norm are poor.

## Conclusion

Our exploration of an effects-based approach strongly suggests that a US doctrine for cyber war needs to understand and incorporate the focal points, and the related saliencies, that are likely to influence how adversaries would interpret a US attack. The basic, effects-based approach does provide a useful starting point. It makes clear why we should not assume that cyberattacks must be deterred by and responded to with cyber means. And some of its more specific findings remain unchanged by the introduction of focal points and saliencies. However, for a variety of situations and types of attacks, including saliences generates significant divergences from the basic effects-based approach. In broad terms, the impact of including saliences in our analysis is to reduce the role of kinetic retaliation in the US cyberwar doctrine. A next step in advancing this analysis is to ask whether we have identified the key possible saliences in cyberwar, and to explore how widely and deeply they are held by individuals and states' decision-makers. Because we are so early in the era of cyberwar, the beliefs and understandings are likely to be weakly formed and to evolve with future experience. Related, because we are in a formative stage, US policies may have the potential to influence the development of certain saliencies; others are likely to stand quite separate from the US policy.

Whether norms have the potential to significantly shape US cyberwar doctrine is less clear. Nevertheless, a norm against cyberattacks against critical infrastructure deserves attention due to the possibility they would result in potentially catastrophic damage. In contrast, a norm against cyberattacks against nuclear C2 appears both infeasible and, even if achieved, too likely to be ineffective to place any hope in. An arms control agreement designed to prevent intrusion into nuclear command and control systems appears more promising. Clearly, the USA will need to employ a diverse range of policy tools in response to the spectrum of cyber threats it faces.

## Acknowledgements

For helpful comments and suggestions, we thank Eric Gartzke, Herb Lin, Scott Sagan and the participants at the Workshop on Strategic Dimensions of Offensive Cyber Operations, Stanford Cyber Policy Program, March 3-4, 2016.

## References

1. Department of Defense. *The Department of Defense Cyber Strategy*. April 2015.
2. Lin H. Escalation dynamics and conflict termination in cyberspace. *Strategic Studies Quarterly* 2012; **6**:46–70.
3. Libicki MC. *Crisis and Escalation in Cyberspace*. Santa Monica, CA: RAND, 2012.
4. Glaser CL. Deterrence of cyberattacks and U.S. national security. Report CW-CSPRI-2011-5. 1 July 2011.
5. Valeriano B, Maness RC. *Cyber War versus Cyber Realities: Cyber Conflict in the International System*. Washington DC: Oxford University Press, 2015.
6. Koh H. International law in cyberspace. Remarks made at Fort Meade, Maryland. September 18, 2012.
7. Snyder GH. *Deterrence and Defense*. Princeton: Princeton University Press, 1961.
8. Utgoff VA. Nuclear weapons and the deterrence of biological and chemical warfare. Occasional Paper No. 36. Henry L. Stimson Center, October 1997.
9. Owens WA, Dam KW, Lin HS. eds. *Technology, Policy, Law, and Ethics Regarding U.S. Acquisition and Use of Cyberattack Capabilities*. Washington DC: National Academies Press, 2009.
10. Brito J, Watkins T. Loving the cyber bomb: the dangers of threat inflation in cyber security policy. *HLS National Security J* 2011; 40–84.
11. Gartzke E. The myth of cyberwar: bringing war in cyberspace back down to Earth. *Intl Security* 2013; **38**:2.
12. Segal A. *The Hacked World Order: How Nations Fight, Trade, Maneuver, and Manipulate in the Digital Age*. New York: PublicAffairs, 2016.
13. Libicki M. *Cyberdeterrence and Cyberwar*. Santa Monica: RAND, 2009.
14. Kushner D. The Real Story of Stuxnet. *IEEE Spectrum*. February 26, 2013.
15. Ackerman S. Snowden: NSA accidentally caused Syria's internet blackout in 2012. *The Guardian*, 13 August 2014.

16. Sanger D, Mazetti MUS. Had cyberattack plan if Iran nuclear dispute led to conflict. *The New York Times*, 16 February 2016.
17. Schelling TC. *The Strategy of Conflict*. Cambridge MA: Harvard University Press, 1960.
18. Schelling TC. *Arms and Influence*. New Haven: Yale University Press, 1966.
19. Clarke RA, Knake RK. *Cyber War: The Next Threat to National Security and What To Do About It*. New York: HarperCollins, 2010.
20. Kreps DM. Corporate culture and economic theory. In: Alt JE, Shepsle KA. (eds), *Perspectives on Positive Political Economy*. New York: Cambridge University Press, 1990.
21. Goldstein J, Keohane RO. Ideas and foreign policy: an analytical framework. In: Goldstein J, Keohane RO. (eds), *Ideas and Foreign Policy: Beliefs, Institutions, and Political Change*. Ithaca NY: Cornell University Press 1993, 3–30.
22. Sugden R. A theory of focal points. *Econ J* 1995; **105**:533–550.
23. Schelling TC. An astonishing sixty years: the legacy of Hiroshima. Noble Prize Lecture. 8 December 2005.
24. https://www.justice.gov/opa/pr/us-charges-five-chinese-military-hackers-cyber-espionage-against-us-corporations-and-labor (29 November 2016, date last accessed).
25. http://www.bloomberg.com/politics/articles/2015-03-17/north-korea-web-outage-was-response-to-sony-hack-lawmaker-says (29 November 2016, date last accessed).
26. Goldsmith J. The DNC hack and (the lack of) deterrence. *Lawfare*, **9** October 2016.
27. Elster J. *The Cement of Society: A Survey of Social Order*. New York: Cambridge University Press, 1989.
28. Tannenwald N. *The Nuclear Taboo: The United States and the Non-Use of Nuclear Weapons Since 1945*. New York: Cambridge University Press, 2007.
29. Goldsmith J. What explains the US-China cyber 'Agreement'? *Lawfare*, **13** September 2015.
30. White House. FACT SHEET: President Xi Jinping's State Visit to the United States. 25 September 2015.
31. Finnemore M, Hollis B. Constructing norms of global cybersecurity. *The American Journal of International Law* 2006; **110**:425–479.
32. Assistant Attorney General for National Security John P. Carlin Delivers Remarks on the National Security Cyber Threat at Harvard Law School, 3 December 2015, available at https://www.justice.gov/opa/speech/assistant-attorney-general-national-security-john-p-carlin-delivers-remarks-national (10 August 2016, date last accessed).
33. Baker S. Steptoe cyberlaw podcast, episode #82: an interview with Jim Lewis. available at https://www.lawfareblog.com/steptoe-cyberlaw-podcast-episode-82-interview-jim-lewis (10 August 2016, date last accessed).
34. Goldsmith JUS. Attribution of China's cyber-theft aids Xi's centralization and anti-corruption efforts. *Lawfare*, **21** June 2016.
35. Sanger D. Chinese curb cyberattacks on U.S. interests, Report Finds. *New York Times*, **20** June 2016.
36. Korzak E. International law and the UN GGE Report on information security. *Just Security*, **2** December 2015.
37. Danzig RJ. *Surviving on a Diet of Poisoned Fruit: Reducing the National Security Risks of America's Cyber Dependencies*. Center for a New American Security. July 2014.
38. Lin H. A virtual necessity: some modest steps toward greater cybersecurity. *Bulletin of the Atomic Scientists* 2012; **68**.
39. U.N. Group of Governmental Experts, Report of the Group of Governmental Experts on Developments in the Field of Information and Telecommunication in the Context of International Security, A/70/174. July 2015.
40. Alexander B. *Downes, Targeting Civilians in War*. Ithaca: Cornell University Press, 2008.
41. Farrell H. *Promoting Norms in Cyberspace*. Council on Foreign Relations Cyber-Brief. 2015.
42. Carter AB, Steinbruner JD, Zraket CA. (eds), *Managing Nuclear Operations*. Washington DC: Brookings Institution Press, 1987.
43. Glaser CL, Fetter S. Should the United States reject MAD? Damage limitation and U.S. nuclear strategy toward China. *International Security* 2016; **41**.
44. Williams RD. (Spy) game change: cyber networks, intelligence collection, and covert action. *George Washington L Rev* 2011; **79**:1162–200.