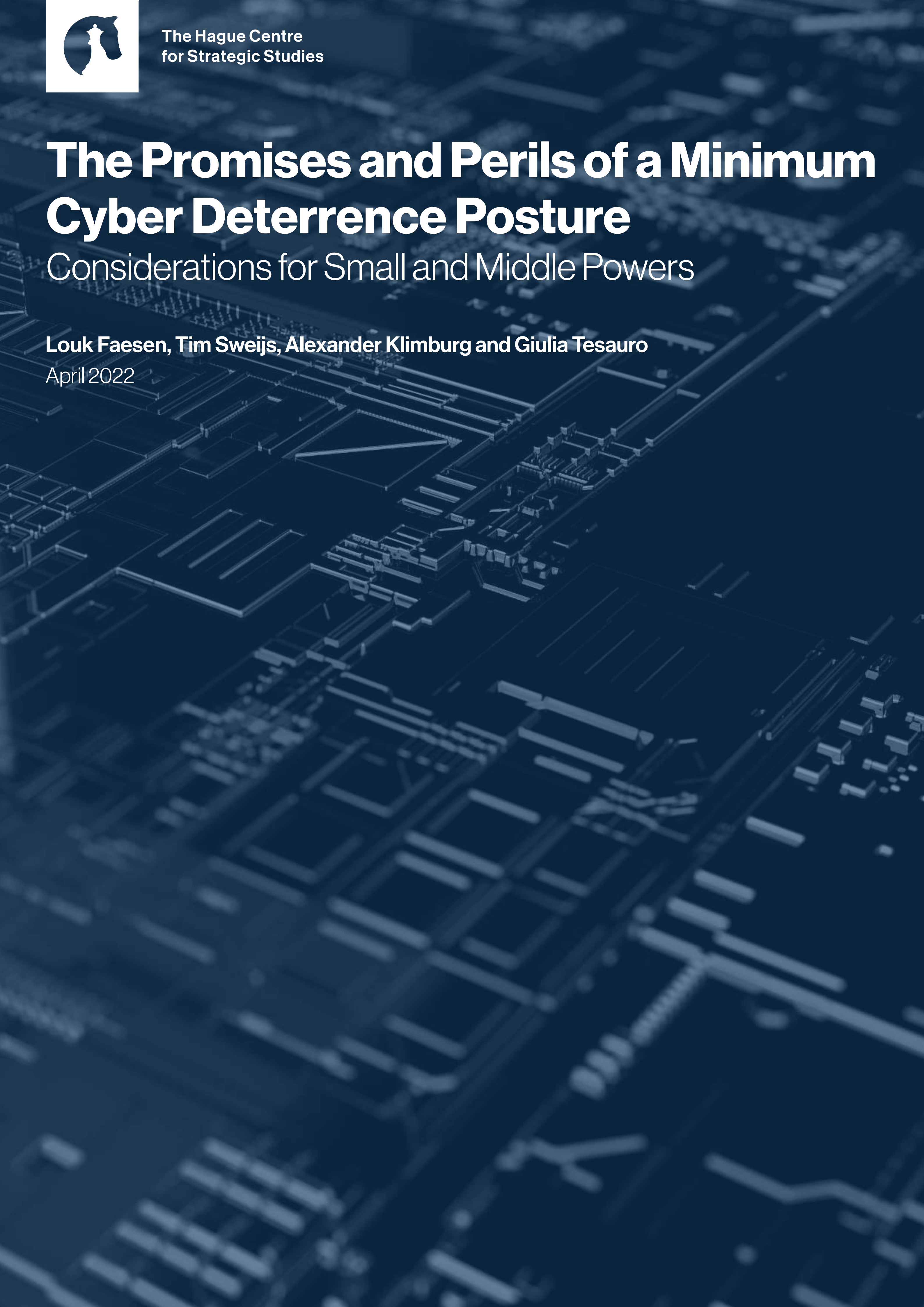
The Hague Centre
for Strategic Studies



The Promises and Perils of a Minimum Cyber Deterrence Posture

Considerations for Small and Middle Powers

Louk Faesen, Tim Sweijs, Alexander Klimburg and Giulia Tesauro

April 2022

The Hague Centre
for Strategic Studies

The Promises and Perils of a Minimum Cyber Deterrence Posture

Considerations for Small and Middle Powers

Authors:
Louk Faesen, Tim Sweijs, Alexander Klimburg and Giulia Tesauro

Introductions by:
Martin Libicki, Michael Daniel, Herbert Lin and Erica Lonergan

Reviewed by:
Pieter Bindt and Frank Bekkers

ISBN/EAN: 9789492102997
April 2022

The research for and production of this report has been conducted within the PROGRESS research framework agreement. Responsibility for the contents and for the opinions expressed, rests solely with the authors and does not constitute, nor should be construed as, an endorsement by the Netherlands Ministries of Foreign Affairs and Defense.

Table of Contents

Executive summary

The emergence of cyberspace provides digitally advanced small and middle powers (SMPs) with a strategic weapons capability that historically has been beyond their reach. Cyber deterrence is understood in this report as deterrence through cyberspace, or, in other words, using cyber means to influence the cost/benefit calculus of the opponent in all domains, not just cyberspace. It thereby focuses on the punishment dimension of deterrence. With the necessary investments and conditions, cyber can be an indispensable tenet of the defensive posture of SMPs and offers an unprecedented punishment capability. It can furnish SMPs with a minimum deterrence capability. The value of minimum deterrence is not to win the war, but to raise the perception that one can inflict unacceptable costs on another actor. Thucydides' famous observation that the strong do what they can while the weak will suffer what they must, may therefore no longer unequivocally hold in today's world.

The development of a strategic weapons capability through cyberspace is not without risk, however: it means that SMPs enter the league of major powers, becoming "medium cyber powers" themselves, thereby putting a target on their back as their strategic significance increases. Although some of these nations may still be unaware of their innate capability to engage in strategic retaliation, other states may already presume that SMPs with advanced economies have the offensive cyber capabilities to engage in deterrence, even if their strategy is to hide those capabilities. Whether they like it or not, cyberspace has changed the deterrence ability of SMPs, just as it has changed the overall conflict landscape.

The road towards a cyber deterrence posture is not an easy one. It poses broader organizational and cultural ramifications that first require concepts and capabilities at the strategic, operational and tactical levels; second, a legal framework; and third, organizational structures and the development of a strategic culture that underlies it. Above all, it requires a strong cyber defense to ward off potential aggressors. This explains the rationale behind a minimum cyber deterrence for SMPs, focusing on the punishment capability, identifies opportunities and challenges, warns for risks, and outlines the requirements for SMPs considering going down this road.

A minimum cyber deterrence capability for small and middle powers

The cyber punishment potential of SMPs may be orders of magnitude less than that of the US or its near-peer cyber powers, but some of these SMPs still possess a *minimum deterrence* capability that, much like the small nuclear arsenals of France, the UK and China, could inflict an unacceptable level of retaliatory punishment on a potential aggressor, despite their overwhelming technical superiority. A minimum cyber deterrence capability can potentially play a critical role in dissuading attacks on the critical infrastructures of SMPs or actions around or above the war threshold that cause large-scale damage. What constitutes unacceptable damage depends on the political objectives and 'nightmare scenarios' of the target. Some countries may find temporary shutdowns of critical infrastructure unacceptable, while others may consider their own regime security as sacrosanct.

In today's geopolitical context, the threat is no longer exclusively military. Interstate relations have become much more hostile and the range of threats governments face have become all-encompassing, fall into blurred areas, and are not just defined by geography. Alliance structures are more complex and arguably do not guarantee the same level of protection anymore as interstate competition deliberately seeks to avoid thresholds that would trigger a collective response. Due to the lack of transparency, states are left guessing as to the overall offensive cyber capabilities of their opponents and allies. In this volatile environment, SMPs need to employ all forms of deterrence, not consider them as individual stand-alone components but as complementary approaches that collectively affect the costs/benefit calculus of adversaries.

Previously confined to deterrence by resilience, by norms, and by entanglement, alongside extended retaliatory arrangements of their powerful allies, cyberspace now offers SMPs an unprecedented form of retaliation of their own.

Previously confined to deterrence by resilience, by norms, and by entanglement, alongside extended retaliatory arrangements of their powerful allies, cyberspace now offers SMPs an unprecedented form of retaliation of their own. The development of offensive cyber capabilities is accelerated by the more assertive US posture in the form of 'defend forward' and 'persistent engagement' that emphasizes so-called in-band responses within the same domain. This posture not only encourages the development of offensive capabilities but also their use. After all, the doctrine postulates that a cyber contest is to be engaged in for it to be understood.

Because of these trends, and perhaps paradoxically, SMPs might do quite well in a 'force-on-force' cyber conflict with a major cyber power. In national cybersecurity, bigger is not necessarily better. Large powers have typically more critical infrastructure parties to coordinate with, more layers of governance that slow down response, and more attack surfaces to cover. The opposite is also true: SMPs on balance are likely to have a relatively easier task in building national cybersecurity because they may be less vulnerable due to their smaller attack surface and fewer parties to coordinate with. This suggests that smaller advanced cyber nations have an inherent advantage in pursuing national cybersecurity compared to their larger counterparts, such as the US, Russia or China, assuming similar levels of digitalisation and dependence on cyber space. The ability of many SMPs to potentially absorb cyber blows, again on balance, is no less significant in a wartime scenario than an advantageous defensive geography would be in a land operation. These nations therefore may possess something which was previously unavailable to them: not just a strategic-weapons capability – a virtual strike force no less potent than a wing of bombers or ballistic missiles – but also a defensive advantage towards larger foes. The sheer number of nations that may be able to compete with and reciprocally threaten a major power could be historically unprecedented. Yet the possible 'high ground' advantage of smaller states in presenting a smaller attack surface in the cyber domain may not have been adequately factored into conventional analysts' calculations of cyber power and deterrence.

A smaller nation's ability to conduct deterrence by punishment is not well understood. From their perspective, it is necessary to understand how to integrate this asymmetric advantage of cyber deterrence into a broader operational framework and project it into their deterrence posture in order to minimize risks. This report focuses on deterrence by punishment, whereas the defensive advantages of SMPs, which are an important prerequisite for the development of such a posture, will be further examined elsewhere.

The unique nature of cyber deterrence

Cyber deterrence is complex because of the multitude of actors, the lack of clear thresholds, and the broader range of second and third-order effects that generate an assortment of risks, which in turn lead to different demands for effective deterrence. First, deterrence can be done by and against a large number of actors. States are not monolithic entities as many departments engage in cyber operations and maintain various degrees of relations with non-state actors (delegation, orchestration, or sanctioning) often leading to a cacophony of actions. This wide range of governmental and non-state actors needs to be influenced and deterred through different means and ends, which requires a thorough understanding of their intentions, capabilities, strengths, and weaknesses, and a differentiated approach to target them.

Second, the thresholds in cyberspace are not as clearly established as in other areas. Cyberspace has shortened the distance between the rungs and thresholds of escalation, and has made potential misperception more likely. If actors have different understandings of thresholds, escalation can spiral without any party deliberately crossing a threshold. Both sides may believe they have not crossed a threshold and are merely responding to the other, leading to a situation where a cyber operation starts in peacetime but ends in wartime through mutual misperception. This is avoided when there are commonly understood thresholds that actors have to cross deliberately. While thresholds can be clarified by publishing national doctrines or statements, by agreeing and implementing norms of responsible behavior, and by acknowledging a country's actions and responses, only a few countries have done so. Although ambiguity within deterrence can be useful to preserve a certain degree of freedom of action, too much ambiguity raises misperceptions and the risk of escalation, which is detrimental to deterrence's utility. States have neither been very successful in establishing thresholds nor at linking specific retaliatory measures to those thresholds. They are therefore encouraged to tie deterrence efforts to more clearly demarcated thresholds while allowing for strategic ambiguity to guard against hybrid tactics that deliberately seek to test such thresholds.

Third, cyber operations create new risks of miscalculations and misperceptions, and thus escalation. The assertion that cyber operations have so far hardly, if ever, escalated into armed conflict is correct albeit with one exception. But viewed within the broader domain of contemporary international relations there are multiple outlets for escalation, in forms short of war. In fact, states are increasingly using cyber, hybrid, and non-military means of power below the threshold of conflict. Overt cyber operations, in particular, risk setting dangerous precedents that change the rules of the game and create a self-reinforcing spiral of tit-for-tat escalations. SMPs should therefore be cognizant of the precedents and escalation dynamics following their actions and to determine their objectives; whether they want to do full spectrum deterrence, including lower level, of intrusions or pursue a minimum deterrence that seeks to deter a specific subset of particularly damaging actions such as attacks on the homeland.

Deterrence through cyberspace: the retaliation options

Offensive cyber can both be very expensive and very cheap at the same time. As a general rule of thumb, the level of preparation for a cyber operation is directly proportional to the level of distinction and limitation of second-order effects. At the highest levels, cyber campaigns, like Stuxnet or Flame, can consume many tens of millions of euros or more to prepare, involve thousands of manhours of cutting-edge bespoke coding, in-depth reconnaissance, testing (even using replicas of physical equipment or networks), and not least the use of a number of zero-day (previously unknown) exploits. Furthermore, they can require special means of inserting the code into the targeted networks, for instance by human agents, which has its own considerable costs attached. A large part of these costs is also reoccurring. As networks change and programs get patched, it is often necessary to repeat most components of the attack chain. Depending on the target, this could easily occur many times a year. Doing cyber legally, in full accordance with international law and the law of armed conflict is much more expensive and time-consuming, but because of its sophistication it also leads to a higher class of offensive cyber capability that is more targeted.

The level of preparation for a cyber operation is directly proportional to the level of distinction and limitation of second-order effects.

However, the vast majority of cyberattacks, like for example Shamoon used to destroy tens of thousands of Saudi Aramco's workstations in 2012, can be much less advanced and still be very disruptive or even destructive– if not at the same targets. Only the most indiscriminate operations can be expected to be launched with limited preparation and still exert a minimum deterrence effect. To this day Iran can, for example, claim to possess a sort of minimum cyber deterrence capability for just a fraction of the cost of the US.

Retaliatory action through cyberspace is separated into two categories of effects: *special cyber operations,* done covertly by intelligence agencies, and *strategic cyber operations,* done overtly by the military. The latter represents a whole new instrument for most SMPs. This subdivision is often based on the legal authorities they depend on and has bearings on international law. The former is less constrained by the bounds of international law and given their covert nature, there are fewer concerns over the precedents of certain actions. Beyond the legal mandates, the main difference is the combination of the target, the effects, and the overall conflict context. For example, a one-off operation in relative peace below the armed attack threshold is more likely to be considered an intelligence special operation. A multi-pronged offensive campaign with cumulative effects that exceeds the armed attack threshold or takes place in a state of belligerency should be considered a regular operation of the armed forces and therefore, if it results in significant damage, a strategic cyberattack. While covert responses only have a direct deterrent effect on the target, compared to the wider deterrence effect of overt measures, the concept of 'special operations' allows more offensive activity to take place, with fewer escalation risks and without officially condoning it, thereby reducing the risk of setting precedents and third-order effects. Therefore, wherever possible, cyberattack options should be done covertly rather than overtly. There are exceptions to the rule of 'special' before 'strategic', based on the scale, impact, and context.

SMPs may lack the resources of the large cyber powers to achieve timely retaliatory effects across a wide range of *countervalue* (wider societal and economic) and *counterforce* (military) sectors. Depending on the strength and robustness of the target's defenses, they will have to concentrate their resources on fewer countervalue sectors rather than distributing them over a large number of sectors. These are likely to have poorer defenses than counterforce targets. Crippling one critical sector will also have a greater effect on an adversary than modest damage across many sectors. This conserves intelligence resources, which are significantly more limited for SMPs, but has significant legal implications that SMPs need to address.

Beyond the counterforce and countervalue targets, a third category that may allow SMPs to achieve an efficient covert cost imposition in peacetime is *counterpolitical* targets. In a similar vein to some sanction regimes, these narrowly-defined targets hold high intrinsic and psychologic value within a country, even if the value is completely hidden from public view and not widely shared. The success of an attack is still likely to depend on the length and level of preparation required, which often involves advance planning and pre-deployment. An approach that focuses on counterpolitical targets accepts the information warfare narrative that is being pushed by some states and which undergirds their overall political objectives and strategy, but refuses its means – namely, subverting the free press and systemic media ecosystem with disinformation.

The intelligence requirements for retaliatory action in cyberspace should extend beyond the standard operational approach and include psychological, political, economic, and other considerations. To fill this gap, the counterinsurgency ASCOPE-PMESII[1] framework can function as a useful tool to determine the broader intelligence needs required for a better understanding of the wider operational environment and the vulnerabilities of the target. What constitutes unacceptable damage depends on the political objectives and 'nightmare scenarios' of the target – some countries may find temporary shutdowns of critical infrastructure unacceptable, while others may consider their own regime security as sacrosanct.

Crafting a comprehensive approach to minimum cyber deterrence

Cyber deterrence is too multifaceted an issue to be dealt with by just the military. To encompass all the various stakeholders, three different approaches are required: a whole of government approach to facilitate *coordination* among government agencies; a whole of nation approach to facilitate *cooperation* between national state and non-state actors, their civil society and industry partners; and a whole of system approach to facilitate *collaboration* with a wide range of state and non-state partners at the international or regional level, most notably within the NATO (whole of Alliance) and EU (whole of Union) context.

The fact that multiple NATO members now possess considerable offensive cyber capabilities adds to the overall credibility of the alliance's deterrence posture. A variety of challenges persist, however, requiring tight coordination of cyber fires and equities (i.e. the wider operations stakes and strategic interests), which need to be deconflicted at speed. The top-tiered cyber allies may also consider different kinds of cyber attack operations that potentially devalue or even invalidate efforts of lower-tiered cyber allies. In case of escalation, allies may drag others into a wider cyber conflict where communication lines between allies and adversaries are likely degraded or damaged, making war termination efforts more complicated. SMPs that are part of an alliance will also face a strategic dilemma when they have to choose between developing their own cyber capabilities to impose unacceptable costs against the adversary or building up the capabilities of the alliance to increase their leverage towards other allies.

[1] The (ASCOPE) PMESII matrix derives from the counter-insurgency context as a basis for targeting in any concept of operations. It contributes to a holistic understanding of the operational environment of friendly, neutral, and threat political military, economic, social information, and infrastructure (PMESII) systems – "a set of interrelated operational variables that provides counterinsurgents with a method to analyze the operational environment through specific filters".

For states without past experience in a whole of government approach to deterrence, the entire concept can be daunting. To advance a whole of government approach, the introduction of a national security council type structure would be a major contribution. Drawing on the results of informal meetings that are already taking place but lack a clear constitutional mandate, a national security council type structure would lead to a better alignment of deterrence efforts within government, as well as domestic and foreign security policy more broadly. The whole of government is the predominant approach in cyber deterrence. It can be facilitated by the whole of nation and -system approaches, which help inform governments, provide additional channels of communication, and contribute to a strategic culture within the government. Given limited resources, SMPs also benefit from leveraging their wider national non-state assets. To this end, liberal democracies are largely limited in coercing or coopting cooperation from non-state actors – industry to a certain extent but especially civil society. Instead, success will largely depend on their ability to convince non-state actors.

Developing a minimum cyber deterrence posture

A national security council type structure would lead to a better alignment of deterrence efforts within government, as well as domestic and foreign security policy more broadly.

SMPs that consider developing a minimum cyber deterrence strategy are encouraged to take the following criteria into consideration for their punishment capability. A prerequisite to such engagement is a strong national cyber defense and resilience. The criteria touch upon attribution, interoperability and common definitions, narrative control, intelligence capability, mandate and legality, the strategic cyber weapon development, coordination with partners and allies, and the management of collateral effects.

1. **Align common terminology or at least clarity of communication and general concepts** with allies and partners. Typologies and definitions are perhaps one of the most elusive components of cyber operations. It is much easier and more useful for policymakers to develop effective descriptions outlining general concepts, rather than fixating on tight definitions This is a crucial first step towards interoperability as SMPs' cyber operations will often operate as part of an alliance structure.
2. **Develop and communicate the ability to attribute in a politically meaningful timeframe**, in which SMPs can technically and politically attribute aggressive behavior from a state actor with sufficient confidence levels, alone or as part of a broader alliance effort, and are willing to share intelligence domestically and with allies to legitimize a response.
3. **Define strategic-end goals and prioritize narrative control** over tactical and operational objectives to avoid path-dependency and the bottom-up problem of putting technical feasibility before political desirability. Narrative control starts with the formulation of a strategy and requires synchronized signaling and strategic communication across all branches of government.
4. **Expand intelligence capabilities** that go beyond the minimum requirements of a common operational picture. This requires support from partner collection and a higher level of resources beyond the operational vantage point. SMPs rely on the intelligence of larger allies, who will only share information under certain conditions (e.g. to what extent do their interests align). Intelligence is required to understand the tactical and operational side of the target (their threat, motivations, tactics, techniques and procedures (TTPs), order of battle, armament and supporting infrastructure, their networks) as well as the broader set of parameters to graph their interlocking political system, for instance by using the PMESSI framework. Creating a cyber common operational picture is sometimes incorrectly viewed as the end goal – however this misstates its utility. It not only facilitates the

targeting process, but also helps to communicate the extent of the larger conflict arena to political decision-makers. Perhaps one of the most-needed intelligence reforms is to open the black box that is cyber conflict to decision-makers. This should not only be directed at the executive branch but also the legislative branch, which should, for example, be able to receive confidential briefings on offensive cyber developments, provided they are thoroughly vetted. Within many SMPs such procedures are missing.

5. **Invest in offensive cyber means capabilities**, including the tools and infrastructure for initial intrusion, for moving within the target, and finally for one or more 'weapons' that need to be deployed without inadvertent release occurring. The costs of cyber capabilities increase on a logarithmic scale, starting at a rudimentary level (indiscriminate, off-the-shelf malware with minimum preparation) going up to the top of the pyramid (discriminate, bespoke malware with lengthy preparation). Many SMPs may lack the resources and intelligence of the top-tiered cyber powers to conduct those high-end operations but can likely still establish a minimum deterrent effect through less sophisticated means.
6. **Assess and limit second and third-order effects,** which are much harder to determine given the complex and dual-use nature of the cyber domain. They can take place outside the area of operations, in other domains than cyberspace, halt allied operations, and undermine the security of other (neutral, allied, or civilian) actors. A detailed intelligence assessment of the target system and its relation to other systems can minimize these operational risks. States also need to take the unintended (normative) effects of cyber operations into consideration that – especially if taken overtly – establish dangerous precedents that can undermine their long-term strategic interests.
7. **Evaluate and balance dilemmas in the equities process,** such as the second and third-order effects, how it impacts other – often intelligence – operations with the same target, how long the vulnerability is likely to be available, and whether others are aware of the same vulnerability to name just a few. These assessments are contingent on how decision-makers rank the respective utilities in their equities process while considering the diverse points of view and capabilities of partners and opponents alike, again requiring international coordination.
8. **Strive for interoperability.** SMPs will often operate as part of a multi-agency or alliance structure, which requires them to synchronize, to be informed, to support, and to execute cyber fires with allies. This is facilitated by common terminology, understanding of the resources, capabilities, goals, strategies, doctrines, and ideological context, as well technological interoperability across ICT systems and policies.
9. **Communicate your offensive cyber abilities** without causing mixed signals or unintended escalation. One way this can be accomplished is through cyber exercises, reported on in the media. Another includes demonstration strikes as a means of signaling intent and capabilities. This would signal not only the power of offensive cyber capabilities but also the ability to find vulnerabilities in the other's systems. By clearly linking offensive use to previously established redlines, SMPs can tie deterrence efforts to more clearly demarcated thresholds, thereby using their punishment capability in a way that minimizes the risk of misinterpretation and inadvertent escalation. Furthermore, in a conflict scenario, sometimes discrete punishment (not publicly visible, but attributable to you) may be the best option to signal resolve while at the same time not pushing the adversary into a public relations corner.

Overall, it is recommended that SMPs recognize the strategic implications of the paradigm shift that is caused by the emergence of cyber as a domain of war. Should they choose not to develop a minimum cyber deterrence capability on the basis of the risks outlined in this report, they need to be aware of the fact that other SMPs may not make that same choice.

1. **Introduction**

Cyber deterrence remains a concept that most states struggle with. Technological advances and realities within cyberspace have changed the extent to which states, in particular, small and middle powers (SMPs) organize and exercise their overall deterrence posture.[2] Traditionally, this posture heavily leaned towards resilience and norms, and entanglement to a certain extent. But there are concerns that defensive or denial effects through cyber hygiene and resilience may fall short against major states, or even just a determined and prolonged attacker. Likewise, the norms for responsible state behavior in cyberspace are relatively new and still suffer from ambiguity and enforcement mechanisms. As a result, states are looking for other ways to affect the cost-benefit calculus of their opponent. Within many of the national security and cyber strategies of SMPs, more coercive forms of deterrence (by punishment) have been mostly avoided but are increasingly gaining popularity.

Arguably "cyber weapons" have caused a more disruptive effect on these considerations than nuclear weapons. Within this new reality, we find ourselves in a similar position with cyber as we did with nuclear arms in the 1950s, but the ramifications are very different. At one point in the late 1950s and 1960s, more than 16 nations, including the likes of Switzerland and South Africa, were considering pursuing nuclear armament. To not do so seemed to invite immediate capitulation in the face of a superior foe. The fact that most nations eventually did not do so was largely based on the immense costs associated with developing and delivering such weapons and the fierce anti-proliferation measures imposed by the initial five nuclear powers. The strategic assessment of geographically smaller nations that their limited size meant their vulnerability to nuclear weapons was higher. Yet, none of these limitations apply to cyber weapons. In fact, in each case, the opposite may be true. There is at present not even a basic academic consensus on how cyber deterrence differs or fits into conventional deterrence frameworks.

Through strategic offensive cyber operations small and medium-sized nations can project power over vast distances and threaten to strike the homelands of major military powers.

The emergence of cyberspace provides wealthier SMPs with a strategic weapons capability that historically has always been beyond their reach. Thucydides' observation that the strong do what they can while the weak will suffer what they must, may therefore no longer hold in today's world. Through strategic offensive cyber operations small and medium-sized nations can project power over vast distances and threaten to strike the homelands of major military powers. Cyberspace has thus opened up not only new and dangerous opportunities for offense, it also affords small and medium-sized nations with new opportunities for defense based on deterrence. They can now credibly threaten to impose enormous costs on potential aggressors. It can even be argued that small to medium sized-nations are less vulnerable too because of their smaller attack surfaces and fewer parties to coordinate with compared to larger nations and their corresponding critical infrastructure.[3] This development marks

[2] Small and medium powers are nations that are traditionally smaller in size who, based on their economic and military strength, are not assumed to do the heavy lifting or take global responsibility in the way that great powers are. Yet they have sufficient power and influence to be able to conduct a normative foreign policy in which they can not only promote their self-interest, but also the interest of a larger or even global community. Within the context of cyberspace, this translates to nations that show a medium or high degree of digitalization, medium to very high cyber defense with variations in readiness, and limited to specialized offensive cyber capabilities that span from Class 3 to 5 (or Tier II to V) Annex III.

[3] As argued by Alexander Klimburg, "Mixed Signals: A Flawed Approach to Cyber Deterrence," *Survival, vol. 62* no. 1, (2020), 107-130. https://doi.org/10.1080/00396338.2020.1715071.

a fundamental change in the strategic landscape of the 21st century. Strategic prudence suggests that these nations should reflect on the implications of this development and on that basis consider their options.

Cyber deterrence – understood here as deterrence *through* cyberspace – can be an indispensable tenet of the defensive posture of small and medium-sized nations. Ultimately, a robust cyber deterrence capability can potentially play a critical role in dissuading attacks on their critical infrastructures and even against territorial invasion – both actions above the war threshold. It is also a key element in modern deterrence at large, where conflict is continuous and below the threshold, threats and damage are cumulative, and where the classic "ladder" of deterrence (rising from entanglement and norms, to resilience and punishment) has been transformed into a "vortex". It is no longer sufficient to engage 'only' in norms and policies fostering entanglement. Instead, nations need to engage in all four parts of deterrence, not considering them as individual stand-along components but as complementary approaches that collectively affect the cost and benefit calculus of other actors. At the same time, the development of a strategic weapons capability through cyberspace is not without risk: it means that small and medium-sized nations enter the league of major powers, becoming "medium cyber powers" themselves. This generates political, military, organizational, and cultural challenges.

This report considers the ramifications of the cyber revolution for the deterrence posture of small and medium-sized nations. It proceeds as follows:

- **Chapter 2** offers a short cyber deterrence primer by Martin Libicki, who outlines its main elements. The chapter then surveys how deterrence has changed as a concept for SMPs, and more specifically how cyber deterrence has evolved, distinguishing between entanglement, resilience, norms, and punishment. It explains how the current escalation and deterrence ladder is transformed into a vortex. It describes how the punishment capability complements the other forms of deterrence.
- **Chapter 3**, introduced by Michael Daniel, offers a clear assessment of the role of actors, actions, and thresholds in cyberspace. The chapter proceeds by offering a basic typology of cyber operations, parses the various actors and offensive means and how they relate to intelligence and information warfare operations, and finally describes the nature of thresholds in cyberspace – a cornerstone of any deterrence strategy.
- **Chapter 4,** introduced by Herbert Lin, this chapter turns to retaliation means and minimum deterrence. It transposes parts of the Single Integrated Operational Plan to cyber, offering operational considerations for targeting and intelligence needs, thereby expanding on conventional strategic terminology – counterforce and countervalue targets – with a third category: *counterpolitical* targets, for which lessons learned are drawn from sanction and counterinsurgency contexts.
- **Chapter 5**, introduced by Erica Lonergan, traces the retaliation paths and their demands on cooperation between government, industry and civil society actors. On that basis, the chapter deduces organizational considerations and singles out requirements for SMPs to develop a minimum cyber deterrence posture.
- **The Conclusion** synthesizes the key insights and outlines ten recommendations to policymakers in SMPs in developing their cyber deterrence posture for the 2020s.
- **Annex I** offers a more detailed historical description of how the concept of deterrence has changed.
- **Annex II** provides additional details about the intelligence agencies from the US, UK, CN, RF that are engaged in cyber operations.
- **Annex III** introduces the tiers that distinguish between the offensive capabilities of states.

2. **Cyber deterrence:** Context, Concepts, Evolution

2.1. **Introduction by Martin Libicki**

The purpose of warfare, as Von Clausewitz explained, is to disarm the enemy, making them unable to resist your will. Although there are wars in which enemies are so disarmed as to effectively disappear, more typical is when the enemy agrees to terms on which war would end. Wars use force to modify behavior, in part by altering the material circumstances of each combatant.

> The essential message of deterrence is: if *you* do *this*, then I *will* do *that.*

Deterrence is another form of behavior modification, but one that works prospectively rather than retrospectively. It is the imposed threat by one side to deliver consequences to another, should that other misbehave in ways variously specified or implied. If deterrence works, then such misbehavior does not occur. Because it is rarely obvious that the other side would have behaved differently in the absence of such a threat, it can often be hard to say that deterrence has, in fact, worked. All we can see is when it fails or has not failed yet.

Deterrence can be a highly attractive proposition. It need not involve war (indeed, war often signifies that deterrence failed). Countries who cannot be defeated can, however, be deterred, if the *costs* of retaliation exceed the expected gain from the action contemplated. Deterrence is, or at least was perceived as, an elegant solution to the problems of nuclear war, a form of combat that could be expected to leave their combatants far worse off whether or not they had won.

This attractiveness has impelled some to apply deterrence theory to other difficult forms of warfare, notably those in which the offense appears to have an edge over the defense. Cyberspace operations are regarded as one such form. In 2007, General Cartwright, then the Vice-Chair of the US Joint Chiefs of Staff, testified that in cyberspace, as in all other domains, the best defense was a good offense. What he meant, given that offensive operations against other hackers were yet to be developed, was that the US ability to carry out offensive operations could be wielded in order to persuade other countries not to do likewise to the United States. Hence, behavior modification.

The requirements of behavior modification, however, are not trivial. One might imagine that a capacity to cause pain suffices, and sometimes it does – but in cyberspace, that alone does not appear to have been enough. It takes more.

The essential message of deterrence is: if *you* do *this*, then I *will* do *that.* Each of the four italicized words have a particular significance. “You” refers to attribution. “This” refers to thresholds, or at least whatever behavior is likely to bring retribution. "Will" refers to the confidence that the other side has that you will, in fact, react as you have so threatened. And "that" refers to the capacity to punish. And those four have to be more-or-less satisfied just to get into the deterrence game. Take each in turn.

The attribution test requires the other side to believe that you (the target) are likely to know who misbehaved (in this context, who carried out the cyberattack) with sufficient confidence to enable retaliation. Fifteen years ago, it was conventional wisdom that attribution was near impossible ("On the Internet, no one knows you're a dog."). But Western governments (notably their militaries and law enforcers) persuaded themselves that attribution difficulties were the primary barrier to deterrence – and correspondingly put considerable work into improving their abilities. At this point, attribution is usually good enough to allow retaliation with high confidence. What lags is the ability to explain attribution convincingly to third parties. It is unclear how important such persuasion is: to wit, will governments stay their hand if they know who did it but are finding it tough to convince others? And how long will attribution stay good? Might state hackers, not having seen much in the way of retribution yet (even as private hackers have been prosecuted), improve their anonymity if the threat of retaliation becomes real?

The confidence or credibility test is also difficult, at least for the cyberattacks that have taken place. The United States promised retaliation for North Korea's attack on Sony Corporation, but it is unclear whether retaliation, in fact, took place (it is unclear who DDoS'd North Korea off the Internet shortly thereafter). Similar promises were made after the Russian interference with the US election; the result was limited to diplomatic retaliation and the addition of a few sanctions. Israel appears to have retaliated against Iran twice (once after a cyberattack on Haifa's tunnel system and once after a cyberattack on water facilities). But track records are otherwise rare to nonexistent.

The thresholds test is laced with confusion. Perhaps the first deliberate and direct death from a cyberattack would cross a threshold for retaliation, but cyberattacks, unlike kinetic attacks, can wreak billions in damage and hurt no one. The law (e.g., the US Computer Fraud and Abuse Act) can act as a threshold for punishing individuals but not states, especially when the machinery for legal responses is well-established but the machinery for cyber retaliation is not. Various US attempts to specify a threshold – e.g., interference with critical infrastructure, the top two percent of all cyberattacks – suffer from great ambiguity.

The capacity to punish, conversely, is well-established, But, in contrast to nuclear weapons, which essentially work everywhere mostly the same, the capacity *to be punished* may vary widely. Some have harder all-around defenses. Some are not particularly vulnerable to cyber-attacks because of the state of their development. When Admiral Rogers asked, in 2015, for a greater emphasis on deterrence, he was asking for a greater capacity to punish – which remains the least problematic of the main four elements.

Having the four elements in place, alas, does not guarantee success. A country may face punishment for something it believes it has a sovereign right to do (e.g., espionage for national security purposes) and stubbornly carry on. Or a country may do the calculus and conclude that it is still better off going ahead. If Putin were convinced that his cyber machinations kept Secretary Clinton out of the White House, would there be any punishment sufficient to dissuade a similar performance, but not so high to raise a serious risk of war?

So, has cyber-deterrence failed on its own? Has it failed for not having been seriously tried? Or has it succeeded by keeping the real disasters (e.g., a prolonged power outage from a cyberattack) from having taken place? At this point, the Owl of Minerva does not even have a departure slot.

2.2. **Deterrence**

This report focuses on the least understood and most cyber-relevant components of deterrence – starting with punishment – and how they may (or may not) convey an asymmetric advantage to small and medium sized states. This is not a reflection of the importance of the respective cyber deterrence dimensions, but rather on the definition of cyber deterrence as *deterrence through cyberspace*, rather than deterring cyberattacks through all four deterrence dimensions. It focuses on the purely cyber means through which actors – in particular SMPs - can bolster their overall deterrence posture within the current geopolitical context. This chapter first introduces the general concepts and waves of deterrence before moving on to describing the different forms of cyber deterrence. Part of this is the description of the recent US doctrine of persistent engagement and its potential effects on the deterrence posture of SMPs, as well as the escalation risks associated with it and cyber operations writ large.

Deterrence is a form of behavior modification by one actor on another and can generally be defined as "the practice of discouraging or restraining someone (…) from taking unwanted actions. […] It involves an effort to stop or prevent an action,"[4] usually achieved by "persuading an adversary that prospective costs would outweigh prospective gains"[5]. In a classical understanding of deterrence, this can be done by denying your adversary benefits (e.g. through better resilience or defenses), by raising the prospective costs through the threat of punishment.

Throughout history, deterrence remains largely focused on raising the costs for adversaries, although the scope and means have adapted to respond to the respective security context. For a detailed historical description of deterrence, see Annex I. With its origins in the field of criminology, deterrence theory started receiving ample attention after the Second World War following the invention of the nuclear bomb.[6] Scholars typically tend to categorize the evolution of deterrence theory in four main waves, starting from the end of WWII through the early and later stages of the Cold War, on to the rise of non-state actors, with some suggesting the emergence of a fifth wave in the context of new military domains and non-conventional capabilities (see Table 1).[7]

[4] Michael J. Mazarr, "Understanding deterrence,"*Deterrence in the 21st century—Insights from theory and practice*, eds. Frans Osinga & Tim Sweijs (Berlin: Springer, 2020), 15.

[5] Lawrence Freedman. "Introduction – The evolution of deterrence strategy and research," *Deterrence in the 21st century—Insights from theory and practice*, eds. Frans Osinga & Tim Sweijs (Berlin: Springer, 2020), 5.

[6] Philosopher Jeremy Bentham applied the principle of deterrence to criminals by supposing that they created a rational cost-benefit calculation when planning to commit a crime. See Freedman, "Introduction", 3

[7] Robert Jervis, "Review: Deterrence theory revisited," *World Politics*, *vol 31*, no. 2 (January 1979): 289-324, https://www.jstor.org/stable/2009945; Jeffrey W. Knopf (2010) The Fourth Wave in Deterrence Research, Contemporary Security Policy, 31:1, 1-33, DOI: 10.1080/13523261003640819. For the fifth wave see: Frans Osinga and Tim Sweijs, Deterrence in the 21st Century – Insights from Theory and Practice", Springer (2021).

Table 1. Five Waves of Deterrence theory[8]

Wave	Central Question	Features
The 1st wave (1940s)	What is the effect of the atomic bomb on international stability?	Exploratory analysis; nuclear domain; great power centric; bipolar system; outside of war.
The 2nd wave (1950s-1960s)	How to defend national security, attain limited political objectives but also control the horrors associated with nuclear war?	Deductive analysis; game-theoretic; operational modelling; nuclear and conventional; great power centric; bipolar system; outside of war. status quo and stability-oriented. Mirror imaging and assumption of unitary actor rationality.
The 3d wave (1970s-1980s)	How to strike a proper (effective and affordable) balance between conventional and nuclear forces?	Empirical; psychological and decision-making perspectives; historical case studies; large n approaches; nuclear and conventional domain; great power centric; bipolar system; outside of war.
The 4th wave (1990s-2000s)	How to deter non-state actors and rogue leaders?	Empirical; multidisciplinary; psychology, terrorism studies; historical case studies; conflict domain; non-state actor centric; unipolar system; outside of war; application in peacekeeping context; incorporated in wider debate on coercive diplomacy and the dynamic relationship between deterrence & compellence; deterrence failures; debate on the utility of precision weapons for conventional deterrence; military theorizing on most effective coercive mechanisms in peace operations to deter and if necessary to compel
The 5th wave (2010-onwards)	What does the deterrence of composite challenges look like?	Partly exploratory, partly empirical; strategic studies; multidisciplinary; perceptions and context; insights from criminology, cognitive sciences, and sanctions literature; all domain and cross-domain, civ and mil; all actor centric; multi-polarity; inside and outside of war; non-status quo orientation; impact of novel technologies.

Historically, SMPs have mainly relied on extended deterrence arrangements (through nuclear weapon-based security guarantees of great powers) or deterrence by denial.

Gaining prominence during the Cold War, treatments of deterrence distinguished between two forms of deterrence: deterrence by denial (making an action unlikely or too costly to succeed) and deterrence by punishment (the threat of grave punishment following that action) to dissuade adversaries from executing it. In the context of the bipolar structure of the Cold War between two nuclear armed opponents, most attention has gone to deterrence by punishment. Glenn Snyder notes that deterrence by punishment is "achieved by the threat of applying some sanction.[9] Thomas [10] and Robert Jervis underlines that the deterrer "manipulates threats to harm others in order to coerce them into doing what he desires.[11] In their conception of punishment, authors initially focused almost exclusively on military tools with deterrence being defined as "the discouragement of military aggression by the threat (implicit or explicit) of applying military force in response to aggression.[12] The purely military nature of deterrence was expanded to other non-military domains, such as economic, diplomatic and information retaliation[13] In addition to threats, the incorporation of assurances, in the sense of "if you don't do this, I won't do that", and positive inducements also received greater consideration.[14]

Historically, SMPs have mainly relied on extended deterrence arrangements (through nuclear weapon-based security guarantees of great powers) or deterrence by denial. While at some point many of these nations seriously considered developing their own nuclear weapons, they

[8] Derived from Frans Osinga and Tim Sweijs, Deterrence in the 21st Century – Insights from Theory and Practice", Springer (2021).

[9] Glenn Snyder, "Deterrence and power," *Journal of Conflict Resolution, vol. 4*, no. 2 (June 1960): 163.

[10] Joseph S. Nye, "Deterrence and Dissuasion in Cyberspace," *International Security vol. 41*, no. 3, (2017), 52 https://www.belfercenter.org/publication/deterrence-and-dissuasion-cyberspace).

[11] Jervis, "Deterrence theory revisited," 291. Note that, in his analysis, Jervis employs this deterrence to refer to the general consideration employed during the second wave of deterrence theory.

[12] Snyder, "Deterrence and power," 167; See also: Mazarr, "Understanding deterrence," 18.

[13] Mazarr, "Understanding deterrence" 18.

[14] Ibid, 19.

were ultimately dissuaded by the high costs and external pressure, and decided to rely on the "lower rungs" of the deterrence ladder. In particular, this meant focusing on resilience, building bunkers rather than weapons, advancing norms, and entanglement, while depending on the extended punishment capability provided by the nuclear umbrella of their allies. This meant that they were only able to proactively manage risk on the margins by promoting international norms and through their allies.

2.3. **Cyber Deterrence**

Cyber deterrence gained more prominence following the widescale DDoS attack against Estonia in 2007 and the Stuxnet operation in 2010. The term was mainly used to describe ways – using means or a broader set of means – to raise the costs of adversarial offensive operations, which, compared to conventional domains, are relatively cheaper and easier to develop and deploy than the total sum of necessary defensive measures. Initially, cyber deterrence emphasized denial efforts or the development of better resilience. This was followed by the realization that the mere focus on defensive measures is not sufficient in deterring aggressors. An expansion towards other forms of deterrence – punishment, norms, and entanglement –followed, despite there being little consensus on the definition of cyber deterrence and the extent to which it could be successfully applied to cyberspace.[15] In a review of scholarly work on cyber deterrence, Stefan Soesanto and Max Smeets identify a variety of ways in which cyber deterrence is defined, based on variation in the means and domains of deterrence as well as the types of actions that are being deterred. They write:

> "First, cyber deterrence can refer to the use of (military) cyber means to deter a (military) attack. Second, cyber deterrence can refer to the use of (military) means to deter a (military) cyber-attack. Third, cyber deterrence can refer to the use of (military) cyber means to deter a (military) cyber-attack."[16]

In this report, we rely on the thrust of the first definitions and define cyber deterrence as the use of cyber means to deter opponents from engaging in undesirable behavior. Soesanto and Smeets also distinguish between three groups of views about the efficacy of cyber deterrence: the first group believes cyber deterrence does not have distinctive problems and therefore works or occasionally fails; the second group argues that cyber deterrence is distinct from conventional deterrence due to the unique nature of cyberspace compared to other domains; and the third group believes cyber deterrence is impossible. This report argues that cyber deterrence may well pose distinctive problems – which will be discussed throughout – but that actors can be deterred through the prospect of punishment by cyber means.

In the context of cyber deterrence, the four-pronged distinction of Joseph S. Nye, places four forms of deterrence next to each other: entanglement, norms, denial and punishment. In practice, these have usually been applied according to a logic of an escalation ladder.[17] It starts with *entanglement* or the existence of various interdependences that, in a successful

[15] Most notably, norms (ramped up by the UN GGE since 2013) and punishment (ramped up by the US indictments and sanctions against state actors for hacking since 2014).

[16] Stefan Soesanto and Max Smeets, "Cyber Deterrence: The Past, Present, and Future," in F. Osinga and T. Sweijs (eds.), *NL ARMS Netherlands Annual Review of Military Studies 2020* (2020).

[17] Nye, 'Deterrence and Dissuasion in Cyberspace', 44-71. Nye does not necessarily describe these four forms of deterrence as a vertical escalation ladder, but rather describes them as four horizontal concepts that – taken together – lead to deterrence.

attack, will impose serious costs on the attacker as well as the victim simultaneously.[18] This is followed by the realization of *norms*: the extent to which diplomatic agreements between states, particularly the norms agreed within the United Nations since 2013, encourage responsible behavior. In both cases, norms encourage predictability and signal interstate consensus on what constitutes bad behavior in cyberspace – a yardstick that the international community can use when calling out hostile actors. Next comes *denial*: the degree to which resilience or cybersecurity deters the attacker because the intended effects of an attack are denied. Finally, *punishment*: the extent to which potential countermeasures deter the attacker. Unlike denial and entanglement, some degree of attribution is necessary for norms to work and for credible punishment to be imposed.

Initially, punishment capabilities in response to cyberattacks used to mainly consist of public attribution, prosecuting hackers, and imposing sanctions. These sorts of actions started to be conceptualized as being part of cross-domain deterrence, whereby one responds through a different domain from the one where the threat originally manifested.[19] It covered the wide range of possible punitive responses that can be taken across the DIME spectrum (diplomatic, informational, military, economic) and included both hard power as well as soft power elements. This led to the creation of different cross-domain escalation paths tracing both vertical escalation – by crossing various thresholds of increased violence – and horizontal escalation – by traversing domains.[20]

The notion of a 'vortex' allows for a non-linear visualization of escalation across different domains and their effect on each other.

Whereas Herman Kahn's original escalation path can be envisaged as a ladder with nuclear annihilation as the highest step, alternative models are said to better explain the increasing interlinkage of different domains in general (primarily diplomatic, military, economic, and informational), and specific military domains (conventional, nuclear, cyber, space), as well as different forms of deterrence (norms, entanglement, resilience, and punishment) that characterize today's conflict landscape. The notion of a 'vortex' allows for a non-linear visualization of escalation across different domains and their effect on each other, in which different forms of deterrence are not placed in a hierarchical order but in a multidimensional construct with many more possible scenarios, onramps, offramps, and outcomes. When escalatory ladders overlap (different means produce ends in the same domain), the outcome is that states produce nonlinear responses to actions.[21] Actions that are typically associated with one domain, for example cyber, can have rigorous effects in the conventional, nuclear and spatial domains. Cyberattacks may be both higher or lower on the concern scale compared to conventional attacks depending on the context and usage.[22] The escalation vortex, also provides insight as to why rule-based approaches to managing escalation pertaining to just one domain are not as successful.[23] Instead, the capabilities that actors have to respond to different levels of provocation in each domain can be represented, which can illustrate vulnerabilities or potential opportunities when developing national security strategies to manage

[18] This does not mean that entanglement in and of itself will prevent conflicts or warts, but it would be equally wrong to assume that interdependence is not taken into account by policymakers and thereby does not reduce the probability of conflict. Robert O. Keohane and Joseph S. Nye Jr., "*Power and Interdependence: World Politics in Transition*" (Boston: Little, Brown, 1977).

[19] See James Andrew Lewis, 'Cross-Domain Deterrence and Credible Threats', (2010); Eric Gartzke and John Lindsay, "Cross-Domain Deterrence", Oxford University Press (2019). Tim Sweijs and Samo Zilincik, "Cross Domain Deterrence and Hybrid Conflict", HCSS (2019).

[20] For more information on cross-domain deterrence see: Tim Sweijs and Samuel Zilincik, "The Essence of Cross-Domain Deterrence," in F. Osinga and T. Sweijs (eds.), *NL ARMS Netherlands Annual Review of Military Studies 2020*, (2020)

[21] Vince, "Cross-Domain Deterrence Seminar Summary Notes"

[22] Paul K. Davis, "Deterrence, Influence, Cyber Attack, and Cyberwar", *New York University Journal of International Law and Politics, vol. 47,* no.2 (2014), 354 https://nyujilp.org/wp-content/uploads/2015/11/NYI203.pdf.

[23] Ibid.

escalation.[24] The escalation vortex is especially important for smaller states who have limited (offensive) capabilities to still place themselves in an advantageous position. That is if they efficiently fill the gaps in their deterrence strategy across different domains.

Persistent Engagement and its Implications for SMPs

In the US, the lack of a dedicated cyber deterrence strategy was strongly criticized by Congress in 2014-15.[25] This prompted a discernable shift towards a more forward-leaning posture by the US Cyber Command (USCYBERCOM), based on the presumption that its deterrence posture in this realm did not succeed in effectively deterring the enemy, in part because it was overly passive and because of the lack of credible cost imposition through countermeasures *in* cyberspace.[26] In other words, the credibility of the American threat of punishment was at stake. Along with offering additional effects like slowing, blocking, or disrupting the enemy, a more proactive stance *in* cyberspace in the form of persistent engagement was proposed to correct this shortcoming. This became known as a posture of persistent engagement which involves assertive prepositioning in cyberspace. It essentially means that USCYBERCOM no longer waits to react until the threshold of an attack is reached, but acts proactively and persistently below the threshold to continuously disrupt the enemy. It pursues opponents across all networks and systems, and allows for forwarding or pre-deployment of exploitations with the aim of gathering intelligence of adversarial operations and planning, while also imposing costs to discourage attacks. The focus on so-called in-band (within domain) responses in cyberspace shows the desire to fight fire with fire, as one defense official described it: "the calculus for us here was that you're pushing back in the same way the adversary has for years. It's not escalatory. In fact, we're finally in the game."[27]

> The US may be the only nation openly discussing deterrence with cyber means, but it is certainly not the only nation thinking along these lines.

Persistent engagement not only encourages the development of offensive capabilities, but also their use. After all, the doctrine postulates that a cyber contest is not only to be welcomed, but to be engaged in so as to be understood. This policy may inadvertently escalate the threat of cyber conflict overall, where a number of smaller states are effectively encouraged to acquire and exercise offensive cyber capabilities. The US may be the only nation openly discussing deterrence with cyber means, but it is certainly not the only nation thinking along these lines. Others may, however, have a very different strategic position in cyberspace, including factors such as openness of the society, available legal measures, limited mandates, but also simply the size of the economic and governmental apparatus. In terms of cyber resilience, bigger is not necessarily better. A major reason for the vulnerable state of US cybersecurity has to do with scale: large nations have inherently more attack surface to cover, and the US easily has the greatest attack surface of all. However, the opposite is therefore also true – smaller nations on balance are very likely to have an easier task in building national cybersecurity. They have a smaller attack surface and fewer parties to coordinate with. If this is true, this means that smaller advanced cyber nations have an inherent

[24] Ibid.

[25] Sean Lyngaas, "Intel chiefs say cyber norms, deterrence strategy still elusive," The Business of Federal Technology, (September 10, 2015). https://fcw.com/security/2015/09/intel-chiefs-say-cyber-norms-deterrence-strategy-still-elusive/250838/.

[26] According to USCYBERCOM's General Paul Nakasone, this resulted in America's enemies no longer believing there are consequences to their irresponsible behavior. US Cyber Command, "Achieve and Maintain Cyberspace Superiority," US Command Vision, (Washington, D.C, April 2018), 2. https://www.cybercom.mil/Portals/56/Documents/USCYBERCOM%20Vision%20April%202018.pdf.

[27] Ellen Nakashima, "U.S. Cyber Command operation disrupted Internet access of Russian troll factory on day of 2018 midterms," Washington Post, (February 27, 2019). https://courses.cs.duke.edu//spring20/compsci342/netid/readings/cyber/wapo-cyber-nakashima.pdf.

advantage in pursuing national cybersecurity compared to their larger counterparts, such as the US or indeed Russia and China. All things considered, smaller nations might do quite well in a force-on-force conflict with a major cyber power. The ability of many smaller nations to potentially absorb cyber blows is no less significant in a wartime scenario than an advantageous defensive geography would be in a ground operation. This has significant repercussions for how they design their strategic deterrence posture. Yet, the 'high ground' advantage of smaller states in presenting a smaller attack surface in the cyber domain may not have been adequately factored into conventional analysts' calculations of cyber power and deterrence, including those that informed the drafting of persistent engagement. As a corollary to this, a smaller nation's ability to conduct deterrence by punishment – a strategy that persistent engagement has heavily incentivized them to pursue – is not well understood.

The 'high ground' advantage of smaller states in presenting a smaller attack surface in the cyber domain may not have been adequately factored into conventional analysts' calculations of cyber power and deterrence.

The destructive potential of these small-to-medium cyber countries may be orders of magnitude less than that of the US or a near-peer cyber power, but some of these countries still possess a minimum deterrence capability that, much like the small nuclear arsenals of France, the UK and China, could inflict an unacceptable level of retaliatory punishment on a potential aggressor, no matter their overwhelming technical superiority. These nations may possess something which was previously unavailable to them: not just a strategic weapons capability – a virtual strike force no less potent than a wing of bombers or ballistic missiles – but also a defensive advantage against larger foes. The sheer number of nations that may be able to compete with and reciprocally threaten a major power could be historically unprecedented. From the perspective of these SMPs, understanding how to integrate this asymmetric advantage of cyber deterrence into a broader operational framework and project it into a Whole of Nation/Union/Alliance deterrence posture will be one of the most important questions in contemporary geopolitics.

Risks of Escalation

The adoption of persistent engagement seems to be based on the premise that cyber operations are unlikely to escalate to conventional forms of violence. Richard Harknett and Michael Fischerkeller assert that actions taken in cyberspace carry limited escalation risk—an assumption that is the linchpin of persistent engagement's viability as an alternative to restraint strategies that rely on deterrence and explicit norms. In a similar vein, Jacquelyn Schneider also concluded that while "it is almost impossible to prove assumptions about escalation… big data analyses, survey experiments, and war games suggests that cyberspace operations rarely create incentives for escalation to armed conflict."[28] Still, risks of escalation should not be underestimated either. While cyber operations have seldom been escalatory in a conventional military sense (leading to the use of large-scale military force), they can be considered escalatory if a wider interpretation of the term is adopted rather than just military escalation. Both points are further explored below, starting with the changing nature of escalation before moving on to the various forms of unintended escalation in cyberspace.

[28] Jaquelyn G. Schneider, "Persistent Engagement: Foundation, Evolution and Evaluation of Strategy," Lawfare, (May 20, 2019) https://www.lawfareblog.com/persistent-engagement-foundation-evolution-and-evaluation-strategy.

First, the assertion that cyber operations have so far hardly, if ever, escalated into an armed conflict or a deployment of conventional weapons, with one exception, is correct.[29] But this assumes a one-sided interpretation of escalation, which limits itself to armed conflict. Viewed within the broader domain of diplomacy there can be multiple outlets for escalation. Ignoring this results in a limited interpretation of escalation dynamics within current international relations in the cyber domain. In the existing context of interstate relations – in which states increasingly use cyber, hybrid and non-military means of power below the threshold of conflict – a situation can, and most often does, escalate in ways other than armed conflict. Consider the intensified trade war between the US and China in which cyber-enabled (economic) espionage played an important role, or the US economic sanctions against Russia as a result of hacking and influencing the US election in 2016. The fact that no conventional weapons were used does not mean that the conflict has not deepened. If we all agree that conflict today has taken hybrid forms of various military and non-military instruments of power, why do we cling to a purely military logic of escalation? It is not inconceivable, and indeed quite appropriate in the present era, that escalation should find its way into areas other than the military.

The second reason is the changing environment in which we find ourselves post-Persistent Engagement. Jacquelyn Schneider draws in part on a study by Brandon Valeriano and Benjamin Jensen who analyzed cyber operations from 2000-2016. They confirm that such operations have rarely led to armed conflict until then, but at the same time warn of the increasing escalation risks of the new US course:

> “We demonstrate that, while cyber operations to date have not been escalatory or particularly effective in achieving decisive outcomes, recent policy changes and strategy pronouncements by the Trump administration increase the risk of escalation while doing nothing to make cyber operations more effective. These changes revolve around a dangerous myth: offense is an effective and easy way to stop rival states from hacking America. New policies for authorizing preemptive offensive cyber strategies risk crossing a threshold and changing the rules of the game. […] [These policy changes] risk exacerbating fear in other countries and creating a self-reinforcing spiral of tit-for-tat escalations that risk war even though each actor feels he is acting defensively-or, as it is called in the scholarly literature, a security dilemma.”[30]

The OSCE, like the Dutch International Cyber Strategy, also observes that activities in cyberspace create much space for ambiguity, speculation, and misperceptions.[31] The concern is that miscalculations and misperceptions between states as a result of activities in cyberspace escalate, with serious consequences for citizens as well as the economy, and the potential for increasing political tensions. Analysts have already pointed to many risks that can contribute

[29] One notable exception is the Israeli bombing of the Hamas Cyber Headquarters, which is seen as a direct response to an earlier cyberattack by Hamas (although this attack should also be seen in the context of a longer-lasting and broader armed conflict). For more information, see for example: International cyber law: interactive toolkit contributors, “Israeli attack against Hamas cyber headquarters in Gaza (2019)”, International cyber law: interactive toolkit, (January 7, 2021) https://cyberlaw.ccdcoe.org/wiki/Israeli_attack_against_Hamas_cyber_headquarters_in_Gaza_(2019).

[30] Brandon G. Valeriano and Ben Jenson, “The Myth of the Cyber Offense: The Case for Cyber Restraint,” Cato Institute Policy Analysis No. 862 (2019), 1-7.

[31] OSCE, “OSCE participating States, in landmark decision, agree to expand list of measures to reduce risk of tensions arising from cyber activities”, (2016), https://www.osce.org/cio/226656; Brief van de minister van Buitenlandse Zaken, “Building Digital Bridges. International Cyber Strategy: Towards and integrated international cyber policy,” (12 February 2017), 12, https://www.government.nl/documents/parliamentary-documents/2017/02/12/international-cyber-strategy, “*In cyberspace we appear to be witnessing a growth in distrust and a danger of escalation and miscalculation. Defensive measures taken by one state can be interpreted by other states as a threat*”.

to escalation between states within the cyber domain (high risk of unintended consequences, misperceptions, miscalculations, etc.), and for cyberspace itself (not a defined military domain).

Lin distinguishes between four types of escalation: deliberate, accidental, inadvertent, and catalytic. Deliberate escalation is carried out "with a specific purpose in mind" through various means, for example, to establish thresholds (explained in the next chapter) whereas the other forms are not considered deliberate by both parties and require further explanation.

First, accidental escalation includes unintended (second or third-order) effects, which in the cyber context can have "greater uncertainty of outcome due to a lack of adequate intelligence on various targets when certain kinds of offensive cyber operations are employed."[32] There simply are a higher number of unknown knowns and unknown unknowns that make up the non-military nature of cyberspace and the omni-use character of 'cyber weapons'. Unintended second-order effects include offensive cyber operations shutting down a system, simultaneously halting an intelligence operation from a third party on the same target. Cyber espionage and attack operations can not only affect the target's system, but also those of third parties that rely on the same system with the same vulnerability – NotPetya being an infamous example. The interconnectivity and interdependence of military, civilian, private, and/or corporate networks and systems, increase the chances, risks, and scope of these unintended effects.

Beyond these operational effects, third-order effects may lead to strategic, legal, and political consequences that run contrary to long-term interests of liberal democracies. Klimburg and others have considered that the US persistent engagement doctrine is likely rooted in an argument that some of the activities witnessed are justified under countermeasures. This is a wide-ranging interpretation, especially as some of the activities that were reported to have taken place under the strategy had a 'kinetic equivalent' component, such as the disruption of the Russian troll factory, the Internet Research Agency in 2017. Klimburg contends that this activity may have created a norm for kinetic equivalent activity taking place in response to propaganda operations – a dangerous precedent the US had no interest in setting.[33] It thereby conveyed a public message implying that it is now acceptable to hack what you consider 'fake news' thereby encouraging disputes about 'bad content'. Ultimately, this may lead to the very thing the doctrine was intended to alleviate: the weaponization of information. Furthermore, by openly communicating about their pre-deployment (rather than being caught in the act) in the Russian electrical grid, the US designated critical infrastructure as a viable vector for coercive signaling. Accordingly, the range of acceptable cyber targets expanded to include critical infrastructure, up to the point of threatening 'mutually assured disruption'. Klimburg implies that if such action had taken place under an intelligence mandate (US Code Title 50), rather than that of the military (Title 10), a distinction further explored in the next chapter, there would have been fewer consequences for international law. How both second and third-order effects influence the organizational considerations of a small to medium-sized nation is explored in the last chapter.

Second, inadvertent escalation happens when an actor takes a deliberate action that it does *not* consider to be escalatory, but which is interpreted as such by the other party. As will become clear in the next chapter, this form of escalation is exacerbated by the ambiguity

[32] Herbert Lin, "Escalation Dynamics and Conflict Termination in Cyberspace," *Strategic Studies Quarterly. Cyber Special Edition 6*, no. 3 (2012), 53. https://www.jstor.org/stable/10.2307/26267261.

[33] Alexander Klimburg, "Mixed Signals: A Flawed Approach to Cyber Deterrence,".

of the actors involved and the ambiguity of actions in cyberspace. The former reflects on the wide range of non-state actors involved with varying links to government, whereas the latter points out that intelligence is always a precursor for an attack. For the effective application of deterrence, it is important to realize that the cost-benefit trade-off of the opponent is based on estimates. Those estimates are based in part on perception. For deterrence to function properly, the perception of costs and benefits must be clear to both the attacking and defending parties.[34] For the victim, it is difficult to distinguish between 'normal' espionage activity, which is not prohibited under international law, and the preparation for an attack in cyberspace. Compared to other domains, intentions and perceptions are even less easy to ascertain and influence. The geopolitical situation is also much more contentious than, for example, during the Cold War. The distinction between threatening with and actually wanting to harm was much clearer back then, as were the capacities and intentions, than it is now in cyberspace. The risk of misinterpretation and consequent inadvertent escalation is therefore much greater now. This challenge is also known as a signaling issue: the signals a party sends through their actions are not clear enough and thus open to multiple interpretations.[35] It is for this reason that states do not just communicate through intelligence or military signals, but through naming and shaming, indictments, sanctions, demarches, norms, and other forms. This requires the laborious and tense process of interdepartmental synchronization of signals.

The geopolitical situation is also much more contentious than, for example, during the Cold War. The distinction between threatening with and actually wanting to harm was much clearer back then, as were the capacities and intentions, than it is now in cyberspace.

Different perceptions do not only occur on an operational level, but also on the strategic and political level if one considers the underlying motivations behind the US persistent engagement doctrine. CYBERCOM clearly believed that it has been too passive in the cyber domain while attacks against the US have only increased. General Nakasone describes that since 2013, the threat assessment has changed from espionage to the disruption of US networks, and cites Iranian operations as an example against which persistent engagement is deployed.[36] But, as noted by several authors, such as Klimburg and Healey, there seems to be a blind spot in the correlation between increasing Iranian cyber aggression and the US cyber operations Stuxnet and Flame against Iran, which became publicly known in 2011 and 2012 respectively, Snowden's revelations in 2013, which hinted at the scale of US cyber operations, or with the assumed role of the US in digitally igniting the Arab Spring in 2011.[37] Healey argues that "America's adversaries probably feel quite confident that they are striking back, not first. Put in stark terms, it is hard, in particular, to blame Iran if it felt the need to respond in kind to the Stuxnet attack. Very close US allies, as well as adversaries, felt that US cyber operations, as revealed by Edward Snowden, showed a lack of restraint."[38] Jervis calls this the "Rashomon effect".[39] The term describes a situation where adversaries do not communicate well with each other, leading each actor to view the situation differently, in line with their own worldview and preferences. Healey emphasizes that it is (too) easy to assume that the one who holds a different worldview will act in bad faith or out of hostility.[40]

[34] Robert Jervis '*Perception and Misperception in International Politics*', Princeton University Press (Princeton: 1976): Jervis 'Deterrence and Perception', 3-30.

[35] Herbert Lin, "Escalation Dynamics and Conflict Termination in Cyberspace", *Strategic Studies Quarterly*, (2012), 46-70.; Healey, "The implications of persistent (and permanent) engagement in cyberspace,"; Jason Healey and Stuart Caudill, "Success of Persistent Engagement in Cyberspace," *Strategic Studies Quarterly*, (2020), 9-15.

[36] Paul M. Nakasone, "A Cyber Force for Persistent Operations" *Joint Force Quarterly* (Issue 92, Q1, 2019),https://ufdc.ufl.edu/AA00061587/00092.

[37] Healey, 'The implications of persistent (and permanent) engagement in cyberspace'.

[38] Ibid, 9.

[39] Robert Jervis, "*How Statesmen Think: The Psychology of International Politics*", Princeton University Press, (Princeton: 2017), 6.

[40] Healey, 'The implications of persistent (and permanent) engagement in cyberspace', 9.

Third, catalytic escalation is triggered by third-party provocation, for example through false flag operations, which are relatively easier to undertake in cyberspace than in the brick-and-mortar world due to the inherent anonymity of cyber operations. The false flag operation against TV5 by Russian intelligence operatives masquerading as ISIS serves as an interesting example to support this claim. It also shows that technical attribution of the actual perpetrator followed after a short delay. It should be noted that attribution includes both technical and political components. At the outset, it involves collecting and analyzing evidence from both technical and other intelligence assets. On the basis of the intelligence evaluation, the state will then make the political decision whether or not to communicate – openly or covertly – about the attribution. This strategy is often used to implicitly signal to opponents that one's technical attribution capabilities have improved markedly and that it has the political will to communicate the attribution, diminishing the margin for plausible deniability for the perpetrator as they are no longer invisible.[41]

Despite these escalation risks, CYBERCOM's vision document - which presents itself as "risk-aware" rather than "risk-averse" – lists only two risks under the "risk mitigation" section: the shortage of resources and personnel to combat all actors and the diplomatic risk that enemies "seek to portray our strategy as 'militarizing' the cyberspace domain."[42] While CYBERCOM is likely to be aware of more risks, an underlying assumption of the Americans is that there can be both "American superiority" and "stability" in cyberspace at the same time.[43] But this does not always go together. Superiority often invokes competition. This is the potential positive feedback loop[44] referred to by Herbert Lin and Max Smeets among others.[45] Robert Jervis says about this: "A failure to anticipate positive feedback is one reason why consequences are often unintended."[46]

2.4. **Main Takeaways**

SMPs have traditionally not been able to independently support robust deterrence postures. While some SMPs considered developing their own nuclear deterrent, most of them were ultimately dissuaded by the high costs from doing so and decided to rely on the lower rungs of the deterrence ladder. In particular, this meant focusing on resilience (large-scale civil defense), and advancing norms and entanglement, while depending on the nuclear umbrella of allies for their punishment capabilities.

The application of deterrence theory to cyberspace strategically empowers SMPs who, for the first time, have the offensive and punishment capabilities not only to defend themselves, but also to strategically and efficiently strike a nation back, regardless of its size.

[41] See the guide to cyber attribution specifying general indicators and examples of successful attribution by Office of the Director of National Intelligence, "A Guide to Cyber Attribution", (September 2018), https://www.dni.gov/files/CTIIC/documents/ODNI_A_Guide_to_Cyber_Attribution.pdf,.

[42] US Cyber Command "Achieve and Maintain Cyberspace Superiority".

[43] Jason Healey, "The implications of persistent (and permanent) engagement in cyberspace," *Journal of Cybersecurity vol. 5,* no. 1, (2019), 9
"*Many assumptions, apparently unrecognized, underlie the belief that the USA can have both superiority and overmatch as well as stability. Yet, in a system as complex as the Internet, "we can never merely do one thing".*

[44] Healey: '*the chain of causation [containing] several feedback loops which interact to amplify (...) conflict.*" Healey, "The implications of persistent (and permanent) engagement in cyberspace", 7.

[45] Herbert Lin and Max Smeets, "What is absent from the US cyber command 'vision", Lawfare (3 May 2018) https://www.lawfareblog.com/what-absent-uscyber-command-vision.

[46] Robert Jervis, "*System Effects: Complexity in Political and Social Life*" (Princeton: Princeton University Press, 1998), 165.

This development has been aided by three trends. First, cyber deterrence is able to better draw on all four rungs of the deterrence ladder, increasing the importance of norms and entanglement while also lowering the (relative) importance of punishment. The transmutation of the deterrence 'ladder' to a deterrence 'vortex' implies that to engage effectively in one field, engagement in other fields is required. So in order to do deterrence by norms effectively, one needs to build a meaningful deterrence by punishment capability.[47] Deterrence by punishment, however, works across the DIME spectrum, allowing the deterrer to hit the attacker with both in-band and out-of-band responses that transcend an individual domain, thereby enhancing the costs of aggression. It also moves the discussion of escalation away from its traditionally linear conception to a *vortex,* illustrating how actions usually associated with one domain, can have significant effects on others as well.

Second, the US persistent engagement doctrine implies that cyber superiority can only be achieved through continuous engagement with adversaries. Opening up debates about risks of escalation, the biggest concern of a persistent engagement approach is that cyber activities might lead to miscalculations and misperceptions, and thus to escalation, and that the public-relations components might set dangerous precedents. While cyber operations have not been escalatory in the traditional sense (resulting in armed conflict), they have been a vehicle for interstate competition and conflict. Furthermore, recent US policy changes allowing preemptive offensive cyber strategies risk crossing a threshold and changing the rules of the game, creating a self-reinforcing spiral of tit-for-tat escalations.

While cyber operations have not been escalatory in the traditional sense (resulting in armed conflict), they have been a vehicle for interstate competition and conflict.

Third, persistent engagement encourages other, smaller states to also develop and use their offensive capabilities. This focus on in-band and constant response in cyberspace presents a new threat and opportunity for smaller states to deploy an "asymmetric deterrence" capability vis-à-vis larger states: a minimum deterrence capability in cyber is both much more affordable and much less definitive (less destructive) then its nuclear pendant. At the same time and as of yet untested is the premise that smaller states may equally be better equipped to do deterrence by resilience, as in standard information security a smaller attack surface goes glove in hand with an easier task in defense.

[47] For an analysis on how norms and countermeasures interact in deterring hybrid – including cyber – conflict, see: Louk Faesen et al, "From Blurred Lines to Red Lines. How Countermeasures and Norms Shape Hybrid Conflict," The Hague Centre for Strategic Studies (September 2020).

3. Actions, Actors and Thresholds

3.1. Introduction by Michael Daniel

Implementing effective deterrence policies requires answering certain questions. Among those questions are: What actions are you trying to deter? Who are you trying to deter? When do you want deterrence to occur? Not surprisingly for those familiar with the field, answering these questions for cyberspace deterrence is hard. Parsing actions, identifying actors, and setting thresholds are challenging problems in the cyber domain. Yet, leaders must distinguish between actions, characterize actors, and set thresholds if they want cyberspace deterrence to function properly.

What makes parsing, identification, and threshold setting so problematic in cyberspace? The difficulty stems from the ambiguity, uncertainty, and duality inherent to cyberspace. These characteristics emerge from how cyberspace is currently constructed, and curtailing those features could mean giving up other, highly desirable attributes, such as scalability, resilience, or reliability. In fact, reducing these negative characteristics while maintaining cyberspace's desirable features forms a key conundrum for cyberspace policy in general, but particularly for deterrence policy. This chapter explores the interaction of the three deterrence questions with these challenging characteristics of cyberspace. Understanding these interactions better would improve the effectiveness of deterrence policy, which in turn would enhance the stability of cyberspace and increase the ratio of benefits to costs in the online world.

Leaders must distinguish between actions, characterize actors, and set thresholds if they want cyberspace deterrence to function properly.

In tackling the first question, what actions should be deterred, policymakers immediately encounter the problems of ambiguity, uncertainty, or duality. Not all actions in cyberspace have these characteristics, of course. Using malware to permanently destroy electric grid transformers is not ambiguous, nor could the action serve any purpose other than an attack. Moreover, deterrence for these types of actions already operates in cyberspace. However, such unambiguous, singular actions form only a small fraction of the malicious activity in cyberspace. Most malicious activity in cyberspace has some degree of ambiguity built into it. For example, the misrouting of internet traffic could come from a deliberate hijacking of the Border Gateway Protocol or from the mis-keying of an IP address into a route announcement. Other actions can serve multiple purposes, as the writers point out in this chapter with multiple examples. As a result, discerning the difference between disliked but acceptable behavior (espionage) versus disliked and unacceptable behavior (setting destructive implants) often requires exquisite intelligence from other domains, such as human intelligence. This chapter will examine the various types of actions occurring in cyberspace and how we can think about them differently in order to reduce the ambiguity, uncertainty, and duality associated with them.

The next question, who is being deterred, is not any easier to answer. Effective deterrence policy relies on clearly identifying the target, because different actors are deterred by different actions. What deters a nation-state intelligence service differs significantly from what deters a cybercriminal. Here again, though, the ambiguity, uncertainty, and duality of cyberspace transform a simple conceptual idea into a thorny practical problem. Who are the actors behind a set of cyber actions? Often the evidence is ambiguous, pointing in multiple potential directions. Or the actors have a dual nature, serving as nation-state actors and criminals at the same time, so how should they be treated? Or the forensic evidence leaves considerable uncertainty because the actors took deliberate steps to obscure their identities or tried to make it look as if another actor was responsible. In addition, different cultural points of view also affect viable deterrence policies; actions that give an East Asian pause might not dissuade actors from the United States – and vice versa.

The third question examined in this chapter also proves challenging to answer, because thresholds in cyberspace are difficult to define. Many actions in cyberspace are scalable or reversible in a way that actions in physical space are not. If a military unit blows up a mobile network tower and shuts down communications in a given area, that damage is not reversible in the short term; however, the same result could be achieved through cyberspace in a manner that is completely reversible. Further, using cyber capabilities that action might even be more targeted, only turning off communications for selected entities. For purposes of thresholds, are those actions the same or different? Thinking about escalation in the form of linear ladder rungs does not work well in cyberspace. Instead, we need to think about thresholds along multiple dimensions at once, considering such factors as context, consequences, scale, and reversibility. For example, a ransomware incident could be a petty criminal act or the equivalent of an armed attack – it depends on the context, consequences, and scale of the incident. In effect, the visual metaphor should be more 3D chessboard than ladder, with many more possible scenarios, onramps, offramps, and outcomes.

The difficulties in answering these three key deterrence questions do not mean that no deterrence occurs in cyberspace. The United States, Russia, China, the United Kingdom, Israel, and Iran all possess sufficient technical capabilities that could cause constant disruption and destruction on the Internet. That they choose not to do so demonstrates that some level of deterrence is already at work in cyberspace.

As a result, the problem is not that there is no deterrence, but rather that our current frameworks and policies are insufficient. While a few types of actions are deterred, too much malicious activity occurs in cyberspace, threatening to radically reduce its benefits. Undeterred crime and disruption also inhibit the ability to find additional value in online capabilities. Moreover, the scope, scale, and cadence of malicious cyber activity already run the risk of turning virtual conflicts into physical war. Given the steady increase in malicious activity over time, that risk only grows with each passing day. None of these outcomes are desirable. As a result, developing deterrence frameworks that cover more actions, deal with multi-faceted actors, and provide logical, explainable, and defensible thresholds for counteraction becomes a critical policy endeavor. It's a tall order, but one that we must meet if we want cyberspace to generate more benefits than harms.

3.2. **Actions**

An important part of *deterrence through cyberspace* is a thorough understanding of the cyber means through which we can deter (actions), who we are trying to deter (actors) and when we want deterrence to occur (the thresholds). First, in terms of the actions or means, the lack of clarity on exactly what capabilities exist in cyberspace means that it is very difficult to comprehensively describe the means (delivery systems or weapons) of such capabilities. There has been a debate about the term 'cyber weapons' ever since they have been used, without many conclusive outcomes on the usefulness of the term.[48] At best, a 'cyber weapon' is a weapon system of omni-use technologies that is extremely difficult for another state to verify due to a lack of transparency. As such, states are only left with the ability to presume – basically to guess – the overall capability of another state (albeit at widely varying degrees of detail) without, in most cases, being able to detail the exact order of battle, table of equipment, tactics, techniques and procedures (TTPs) or other basic information – unless the intelligence assessment is very complete. Instead, it makes more sense to approach cyber weapons as operations. A basic typology is used to parse the wide range of cyber operations and the important link and tension they have with intelligence and information warfare operations.

Second, within the conventional and nuclear forms of deterrence, states have been and remain the dominant actor. In cyberspace, things are a bit more complicated. Deterrence is done by and against a larger number of actors. This chapter parses the actors across state and non-state entities and explains how they intersect.

Deterrence is done by and against a larger number of actors.

Third, deterrence implies setting thresholds: attacking *this* would warrant a response or escalation. How successful have we been in linking deterrence efforts to behavior benchmarks? To answer this question, the legal and normative thresholds are used as a starting point for analysis before moving on to actions states take in establishing thresholds.

3.2.1. **A Cyber Operations Typology**

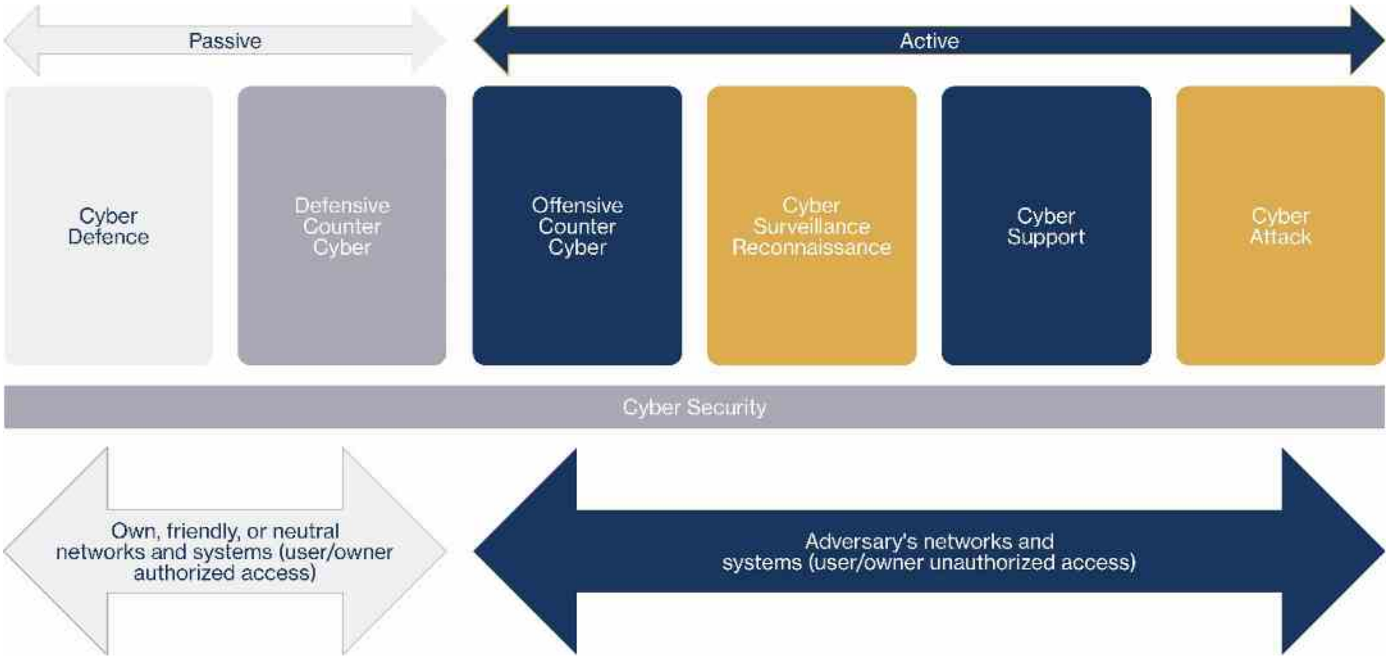
Numerous typologies exist to describe the wide range of cyber operations as nearly every country uses distinct typologies that undergo constant change. The issue is simply that cyber operations can cover the entire gamut of overt and covert action in cyberspace, meaning that virtually nothing is excludable. Traditionally, there is a very wide span of different understandings on how distinct elements of espionage, kinetic-equivalent, and psychological influence operations are categorized in and through cyberspace. There is also a practical differentiation between cyber effects that occur directly in the kinetic battlefield conducted at speed with and against military equipment (which usually are an approximation of Electronic Counter Measures), and strategic cyber, which largely uses conventional Internet technologies or even the Internet itself, and is often marked by a much slower operational tempo in multi-use computer networks (often associated with Advanced Persistent Threats). Although battlefield and strategic cyber may overlap, the former, sometimes called Cyber Electro-Magnetic Activities (CEMA), is outside the scope of this report as it can be considered a continuation of the existing Electronic Warfare practice that militaries have been developing for the past seven decades and therefore, not a paradigm shift in its own right. The cyber operations discussed here are those that can best be summarized as strategic cyber, for even though they may have tactical battlefield relevance, they use the infrastructure of Internet technologies or even the global Internet itself to pursue their mission.

For the purpose of this report, a basic typology is used (see Figure 1). The first distinction can be made between passive and active measures. Passive measures take place in friendly or neutral networks and systems, and largely coincide with what the US describes as blue space and gray space. They are generally understood to be defensive in nature and include cyber

[48] Alexander Klimburg and Louk Faesen, "Balance of Power in Cyberspace," in Dennis Broeders and Bibi van den Berg (ed.), "*Governing Cyberspace: Behavior, Power, and Diplomacy*", (United Kingdom: 2020).

defense operations and to some extent also defensive counter cyber operations. The former consists of purely defensive security measures, such as basic security controls, firewalls, risk and patch management, vulnerability assessments, scanning, and monitoring, that are aimed at enabling freedom of action and force protection. Defensive counter cyber operations go beyond purely preventative measures and are a defensive response to an intrusion or potential threat within the owned networks and systems, or those to which authorization has been granted. They are often aimed at detecting and terminating an attack, mitigating its effects, and recovering from it.

Figure 1. Cyber Operations Typology


Passive
Active
Cyber Defence
Defensive Counter Cyber
Offensive Counter Cyber
Cyber Surveillance Reconnaissance
Cyber Support
Cyber Attack
Cyber Security
Own, friendly, or neutral networks and systems (user/owner authorized access)
Adversary's networks and systems (user/owner unauthorized access)

Active measures take place in the adversary's networks and system and largely coincides with what the US describes as red space. They include cyber ISR (intelligence, surveillance and reconnaissance), offensive counter-cyber operations, cyber support operations and cyberattack operations.

Cyber Intelligence, Surveillance, and Reconnaissance (ISR)

Cyber intelligence, surveillance, and reconnaissance operations (ISR) are effectively covert intelligence operations to gather information from target and adversary systems and networks to support other cyber operations. Cyber operations are tailor-made combinations of intelligence, intrusion, and attack, and it is seldom clear where one phase ends and another begins. However, in most Western typologies a distinction is made between espionage, formerly known as Network Exploitation (CNE), and attacks, formerly known as Computer Network Attack (CNA). This was a useful way to express obvious legal differences between "looking at data" and the "blocking, manipulating or destroying it."[49] A well-known example of this is Title 10 (armed forces) and Title 50 (espionage) authorities, further described below. At the same time, "it obscures the fact that in nearly all cases attacks (CNA) requires espionage (CNE) to be effective – and what is CNE can be switched to CNA with the ease it takes a uniformed soldier and a civilian spy to switch chairs."[50]

What is CNE can be switched to CNA with the ease it takes a uniformed soldier and a civilian spy to switch chairs.

Within the military typologies, a distinction can be made between Cyber ISR and preparation of the environment or battlefield (OPE). The former contributes to the overall situational awareness and understanding of the recognized cyberspace picture within the common operating picture, which informs and enables decision-making for both offensive and defensive operations. Preparation of the battlefield constitutes intelligence operations that are carried out with the intent of planning and preparing for an attack, instead of focusing on intelligence gathering.[51] The difference between such an imminent preparation for attack (e.g. OPE) and 'simple espionage' can be hard to distinguish for the defender, making inadvertent escalation much more likely due to a failure to correctly interpret intent, which is further explored in the section on thresholds.

Upstream, downstream, and endpoint collection

The range of cyber ISR is enormous. Intelligence can include all kinds of basic cyber intelligence gathering (including using open source and commercial means and SIGINT) that can equally depend on large network monitoring and big data mining operations. Many of the national intelligence programs can be broken down into three broad categories: upstream collection, downstream collection, or endpoint collection. Upstream targets the major cables and routers called the Internet backbone, but could also include satellites, while downstream targets large content intermediaries and processors of data like Google or Facebook, much like the NSA's Prism program through which large content intermediaries agreed to automated access of some of their data by the NSA. Endpoint collection is directly targeting the devices of entities.[52] With increasing sophistication not only is the vast majority of collection processes highly automated, but also the 'minimization' and even analysis process itself. Before surveillance starts – when a human analyst looks at the data – an automated minimization process is run to filter out relevant data or data that may be unlawful, so very few bits of data are actually viewed by human analysts.

Within most liberal democratic countries, military and intelligence operations fall under different legal paradigms. Overall, the main difference is that military operations are governed by International Humanitarian Law (IHL) and customary law in wartime, and by the UN Charter and customary law in peacetime. International law is, however, silent on explicitly prohibiting espionage.[53] Within the US, for example, Title 50 Authorities primarily allows for covert intelli-

[49] Klimburg, *The Darkening Web. The War for Cyberspace* (New York: Penguin Putnam, 2018), 152

[50] Ibid.

[51] FM 3-12,"Cyberspace and Electronic Warfare Operations," (Washington DC: April 2017), 1-10

[52] Ibid, 173

[53] Except for some undefined international gentlemen's agreements, espionage is only really regulated on the national level.

gence operations to gain access to hostile targets.[54] Such operations still make up for the bulk of CYBERCOM's operations under persistent engagement, but as soon as they start to affect the target's networks, by disrupting them for example, they move into Title 10 Authorities. These operations fall within the scope of international law and, unlike Title 50 operations, involve more legal concerns, frictions, and consequences.[55] This does not mean there are no restrictions on covert intelligence operations. First, such operations are primarily curbed by national legislation in liberal democracies, while other regimes may be less legalistic and offer more room to maneuver. Second, a more implicit limitation can be identified through international law. Consensus on the legal application of sovereignty and nonintervention to cyberspace and intelligence operations remains debated, but for most European states a restraint is implied on far-reaching covert operations. The use of force by intelligence agencies is a complex and under-discussed legal field.[56] But some experts believe the target itself functions as a distinguishing feature on the nature of an operation – being intelligence or sabotage. During an armed conflict, critical infrastructure may be a legitimate target for sabotage, taking into consideration IHL principles. Infiltrating such networks during peacetime would, however, generate no political intelligence, but would mainly be intended for preparation of the battlefield. Doing passive reconnaissance through SIGINT is accepted, but according to Boeke and Broeders, "leaving implants in the adversary's networks seem as illegitimate as laying remotely controlled sea-mines inside a port in peacetime. In doing so, the agencies step outside their own legal paradigm."[57]

The level of preparation required for a cyberattack is directly proportionate both to the level of target and effect distinction one wants to achieve, as well as the defenses of the target.

This interpretation presents a significant challenge for those nations that follow a restrictive approach to cyber operations. How do you prepare for a kinetic-equivalent cyber conflict, when in most cases this will require some kind of access and preparation on a not-yet-enemy network before the conflict occurs? Unlike for the vast majority of situations in the physical domains (and for which international law was written), it is often not possible to pick up a 'cyber weapon' and launch it, as is, at a known target. Both the attack path (ingress) and the actual 'cyber weapon' will often need to be tailor-made for each target, and in each case updated regularly to suit changes in circumstances, such as patched software or changes in routing. In a workshop conducted for this report, it was noted that "the level of preparation required for a cyberattack is directly proportionate both to the level of target and effect distinction one wants to achieve, as well as the defenses of the target".[58] Put differently, some debilitating attacks can likely be launched with minimum preparation, against a fairly soft target – for instance, Shamoon (the suspected Iranian attack on Saudi ARAMCO) could be such an example. However, with increasing sophistication the time and labor investment rise exponentially. The effort required to launch Stuxnet (a US/Israeli cyber operation against the Iranian nuclear enrichment program) probably required 100 to even 1.000 times the resources (man-hours and direct financial investment) of Shamoon.

[54] See, among others: Max Smeets, "US cyber strategy of persistent engagement & defend forward: implications for the alliance and intelligence collection," *Intelligence and National Security*, (2020), 444-453

[55] For more information on Title 10 and Title 50, see: Andru E. Wal", "Demystifying the Title10-Title50 Debate: Distinguishing Military Operations, Intelligence Activities & Covert Acti"n," *Harvard National Security Journal*, (2011) https://heinonline.org/HOL/P?h=hein.journals/harvardnsj3&i=85.

[56] Sergei Boeke and Dennis Broeders, "The Demilitarisation of Cyber Conflict," *Survival vol 6*, no. 60 (November 2018), 78.

[57] Ibid.

[58] High-level workshop conducted in November 2021 for the purpose of informing this report.

Offensive Counter Cyber Operations

Both defensive and counter cyber operations largely coincide with what JP 3-12, the leading US doctrinal publication on cyberspace operations, describes as countermeasures in cyberspace: "impairment of the operational effectiveness of enemy activity" that can be taken preemptively or reactively.[59] Defensive measures reflect the "internal countermeasures" that are taken within own networks and include activities such as the closing of router ports being used by an adversary for unauthorized access or the blocking of malware that is beaconing out from owned networks. Offensive counter cyber operations are specifically targeted against adversaries' offensive cyber capabilities or aimed at determining the origin of an attack that involves pre-emptive or preventive counter-operations in cyberspace. They largely coincide with external countermeasures that target malicious cyberspace activity which take place outside of the owned networks. These include actions that are usually "nondestructive/ nonlethal" and "minimally intrusive techniques to interdict or mitigate threats" or that "spoof or otherwise negative the effectiveness of adversary sensors or defenses."[60] Their impact or effect is typically limited to adversary malicious activity, not the overall adversarial system, and stop when the threat stops. One example may be the US CYBERCOM operation that temporarily took the Russian Internet Research Agency offline to prevent it from disseminating disinformation to influence the American midterm elections in 2018 (as it did during the preceding presidential elections).

Cyber Support Operations

Cyber support operations are operations in support of other activities. This can include a supportive function for civilian authorities or traditional military operations in the kinetic battlefield, such as the suppression of enemy air defenses. The latter is usually referred to as CEMA operations mentioned previously, such as the *Suter* program that was used in Operation Orchard. Another supportive role includes cyber-enabled influence operations that produce psychological effects on the perception or behavior of the adversary, which are further described under information warfare. While countries like Russia and China engage in this activity during peacetime and at a strategic level, for many of the Western countries, they are mainly restricted to the tactical battlefield. There are nonetheless indications that western intelligence and military agencies also engage in such operations as part of their covert online influencing operations. Examples date back to 1994 when the US engaged in this activity to convince Haitian security to surrender to US forces without a fight, which proves particularly useful against less-developed cyber nations. More recently, the Snowden disclosures offered insight into the UK's covert online influencing operations. The typologies used by liberal democracies to describe these operations have undergone many changes - information warfare, psychological operations, and information operations to name just a few – and are often used interchangeably. Such doctrinal confusion has been described in more detail by one of the authors elsewhere, and more recently by Herb Lin.[61] They are further untangled in the section on cyber operations and information warfare.

[59] US Department of the Army "Joint Publication 3-12: Cyberspace Operations," (8 June 2018), II-7.

[60] Ibid.

[61] Herbert Lin, "Doctrinal Confusion and Cultural Dysfunction in DoD: Regarding Information Operations, Cyber Operations, and Related Concepts;" *Cyberdefensereview* (2020); Klimburg, *The Darkening Web*.

Cyber Attack Operations

Cyber attack operations are a form of *fires*. Unlike intelligence activity, which remains clandestine to be effective, cyberattack operations will often, but not necessarily, be apparent to the system operators, either immediately or eventually, since they affect or remove user functionality.[62] They infringe on the availability and integrity of data and ICT systems to achieve the desired effects (i.e. degrade, disrupt or destroy), or manipulation that leads to denial effects in the physical domains.[63] This can be accomplished through denial of service or destructive malware insertion and other means, such as phishing, and ransomware attacks. The NATO Allied Joint Publication on 'Cyberspace Operations' defines *degrade* as a way to "deny access to, or operation of, an asset to a reduced level of its capacity and/or performance." Disruption can be considered a more extreme case of degradation, where it "completely denies access to, or operation of, an asset for a period of time." Finally, it defines *destroy* as "to completely and irreparably deny access to, or operation of, an asset. The asset is affected to the maximum extent, both in terms of outage time and damage caused."[64]

The exact nature of kinetic-equivalent effects, formerly known as "Computer Network Attack" [65] and now known as "Offensive Cyber Effect Operations"(OCEO)[66] is ubiquitous. In fact, the latter expanded the level of ambiguity even further, making it virtually impossible to know what would be included and excluded from its scope. For instance, OCEO targeted at a power grid could of course mean switching off the grid. But it could also mean destroying the grid to many different degrees, including to the extent that it cannot easily be reconstituted. And finally, it could also mean something completely different where for instance the power grid is simply repositioned to be used as an espionage tool, or even as a weapon itself. This lack of clarity means that it is very difficult to describe comprehensively what the intent and the means (delivery systems or weapons) behind such operations are.

One of the main overarching distinctions made is between cyber and information warfare operations, with NATO allies focusing predominantly on the former, and China and Russia on the latter.

3.2.2. Cyber Operations & Information Warfare

One of the main overarching distinctions made is between cyber and information warfare operations, with NATO allies focusing predominantly on the former, and China and Russia on the latter. This difference can be traced back to the definition of *cybersecurity* or *information security*. The Western interpretation of information security (or infosec) is mostly used by the technical cybersecurity community that relies on the ISO definition that only concerns itself with the status of the data from a technical point of view: "the purpose of information security is to protect and preserve the confidentiality, integrity, and availability of information (CIA triad)."[67] The Chinese and Russian interpretation, by contrast, perceive information security in much broader terms by putting more emphasis on the content of data as a potential threat to

[62] U.S. Department of the Army, "Joint Publication 3-12:Cyberspace Operations," II-7.

[63] Ibid.

[64] Ministry of Defence, "Allied Joint Publication-3.20 Allied Joint Doctrine for Cyberspace Operations," NATO Standardization Office (NSO), (United Kingdom: 2020).

[65] Committee on National Security Systems (CNSS), "Committee on National Security Systems (CNSS) 'Computer Network Attack (I),'" Glossary (Computer Security Resource Center, 2015).

[66] Presidential Policy Directive /PPD-20, "U.S. Cyber Operations Policy," Federation of American Scientists – Intelligence Resource Program.

[67] ISO / IEC, "Information technology – Security techniques – Information security management systems – Overview and vocabulary," ISO/IEC 27000: 2018(E), (Geneva: ISO, 2018).

domestic stability.[68] It can encompass critical or dissenting content that is deemed undesirable by the state. As a result, these actors are mainly preoccupied with information warfare.

This does not mean that liberal democracies are not engaged in online influence operations. Overall, a distinction can be made between information operations, psychological operations, covert influencing and strategic communication

Information operations are described "as having five specific components or dimensions: computer network operations (CNO), psychological operations (PSYOPS)[69], signals (maintaining communication), military deception (MILDEC), and intelligence/counterintelligence."[70] Indeed, the definition of information operations puts an equal emphasis on the cyber component of CNO and the psychological warfare components of PSYOPS and MILDEC. This degree of overlap has produced a level of confusion but also lateral freedom in the conduct of US offensive actions. Information and psychological operations are reserved to the battlefield, a localized military campaign at the tactical or operational level rather than a national campaign or strategic weapon that is directed at the political leadership of another nation.[71]

Covert influencing is purely an intelligence operation that can be carried out in peacetime. In contrast to strategic communication, these operations are usually very targeted at an individual or small group and effectively has no limitations on the way it is employed. The British Operation Cupcake, for example, replaced the content of a Jihadi manual with instructions on how to make a DIY bomb with a cupcake recipe. In the same way, cyberspace changed traditional covert espionage, it has also changed covert influencing in a way that takes "paranoia to a new level."[72] Or so at least the UK's Joint Threat Research Intelligence Group (JTRIG) boasted about its activity. Based on leaked documents, this unit engages in online covert action to collect intelligence, plant propaganda, or to deny, disrupt, degrade, deceive or discredit opponents. They do so through hacking, honey traps, false information, or false flags (posing as an enemy).[73] Their targets extended beyond the usual suspects, such as hostile nations and their leaders, military and intelligence community, and included people suspected (not charged) of ordinary crimes.[74] Operation QUITO is known as the "pioneering effects

[68] Within Russia it has been defined as "a state in which personal interests, society, and the government are protected against the threat of destructive actions and other negative actions in the information space." The Ministry of Foreign Affairs of the Russian Federation, "Convention On International Information Security," (September 22, 2011).

[69] Psychological Operations (PSYOPS) are further defined in JP3-13.2 Their purpose is to "convey messages to selected foreign groups to promote particular themes that result in desired foreign attitudes and behaviors" and "shape the security environment to promote bilateral cooperation, ease tension and deter aggression." United States Joint Forces Development, "Joint Publication 3-13.2: Psychological Operations," (January 7, 2010).

[70] United States Army, "Joint Publication 3-13 Information Operations", (November 27, 2012); Klimburg, *The Darkening Web,* 151.

[71] Klimburg, *The Darkening Web,* 151.

[72] Carl Miller, "Inside the British Army's secret information warfare machine," Wired (November 14, 2010).

[73] An impressive catalog of JTRIG tools and techniques have been leaked and include "the ability to manipulate the results of online polls, artificially inflate pageview counts on web sites, 'amplif[y]' sanctioned messages on YouTube," and plant false Facebook wall posts for entire countries. The Intercept, "JTRIG Tools and Techniques," (July 14, 2014); Matt Kennard and Mart Curtis, "Revealed: Veterans of the UK military's cyber warfare unit are teaching school children how to launch cyber attacks;" markcurtis.info (July 15, 2020).

[74] The Intercept report that "In fact, the discussion of many of these techniques occurs in the context of using them in lieu of 'traditional law enforcement' against people suspected (but not charged or convicted) of ordinary crimes or, more broadly still, 'hacktivism', meaning those who use online protest activity for political ends." Glenn Greenwald, "How Covert Agents Infiltrate the Internet to Manipulate, Deceive, and Destroy Reputations", *The Intercept*, (February 25, 2014).

operation" carried out covertly by JTRIG from at least 2009, using cyber operations and disinformation to prevent Argentina from encroaching on the Falkland Islands.[75]

Strategic communication, or StratCom, aims to influence the hearts and minds of a much wider audience through mass-scale media (television, radio, public Internet forums), but is also limited in that it can only emphasize or de-emphasize certain information.[76] It does not allow for the widespread and deliberate dissemination of lies. Within this context it is worth noting the well-known American anti-propaganda law, the Smith-Mundt Act, which prohibits the US government's propaganda efforts from reaching American citizens.[77] The act does not prohibit the use of propaganda against foreign entities, but it does invoke a more cautious approach to its broadcasting efforts, and it significantly limits them as they may not reach any US citizens. While this act has been subject to many amendments, including one in July 2013[78] that significantly relaxed many of the restrictions (or protections) in it, it is not yet clear to what extent the amendment changed the mode of operation and the scope of the StratCom efforts.

Liberal democracies use information operations predominantly at the tactical level of the battlefield and broader StratCom campaigns in peacetime adhere to a fact-based approach.

The first Tallinn Manual, dealing with cyber warfare in the context of International Humanitarian Law, stipulates that misinformation may be used to mislead adversaries but must distinguish between civilians and combatants[79] and cannot harm the former in pursuit of the latter.[80] The law itself is dubious in applying it to cyberspace, notably in suggesting that "media used for military purposes may be lawfully attacked"[81] but not detailing *how* this distinction is to be made in regard to the complex role of social media platforms as potential dual-use vectors of information operations. Despite this legal ambiguity, liberal democracies use information operations predominantly at the tactical level of the battlefield and broader StratCom campaigns in peacetime adhere to a fact-based approach. This starkly contrasts with the relativist doctrine of Russian information warfare campaigns that utilize a so-called 'plurality of truth' to spread falsehoods as part of broad-scale information warfare both in peacetime and wartime.[82]

Russia's view of the importance of information as a weapon was clarified in the 2016 Information Security Doctrine, in which it distinguished two forms of informational attacks: the technical attack and psychological attack.[83] It is mostly concerned with the latter, and nearly

[75] Operation QUITO was carried out in support of the Foreign Office's goals concerning Argentina and the Falkland Islands. The supporting role resurfaced in other documents, where it was mentioned that "the Foreign Office is looking for advice" for an upcoming visit to Chile to counter a trend of growing support behind Argentina among South American attitudes. Glenn Greenwald, The Intercept, (February 25, 2014).

[76] It is predicated on "efforts to understand and engage key audiences to create, strengthen, or preserve conditions favorable for the advancement of government interests, policies, and objectives through the use of coordinated programs, plans, themes, messages, and products synchronized with the actions of all instruments of national power." U.S. Department of Defense, "Strategic Communication Joint Integrating Concept", (7 October, 2009), B-10: https://www.jcs.mil/Portals/36/Documents/Doctrine/concepts/jic_strategiccommunications.pdf?ver=2017-12-28-162005-353.

[77] Mac Thornberry, "H.R.5736 – Smith Mundt Modernization Act of 2012", House Committee on Foreign Affairs, (May 10, 2012).

[78] Klimburg, *The Darkening Web.*

[79] Determining the legal status of an individual under IHL presents difficulties. Overall, ISIS members who directly participate in hostilities in Syria and Iraq may be lawfully targeted by military operations. Holli Edwards"Does International Law Apply to the Islamic State?", *Geneva Centre for Security Policy*, no.1 (2017); Christophe Paulussen, Hanne Cuyckens, and Katharine Fortin, "The Prosecution of Foreign Fighters Under International Humanitarian Law: Misconceptions and Opportunities", International Centre for Counter-Terrorism, (December 13, 2019).

[80] GroJIL: "The Truth Under Siege: Does International Humanitarian Law Respond Adequately to Information Warfare?", Groningen Journal of International Law (2019).

[81] Michael Schmitt, "Tallinn Manual on the International Law Applicable to Cyber Warfare", NATO (2013).

[82] Stefan Meister, "Understanding Russian Communication Strategy: Case Studies of Serbia and Estonia", Institut für Auslandsbeziehungen, (2018).

[83] Ministry of Foreign Affairs of Russia, "Doctrine of Information Security of the Russian Federation", (2016).

all technical attacks (including cyber and electronic attacks) are coordinated or supplemented with a psychological effect in mind. Violent active measures such as "kidnapping adversary officials" and "physical destruction of adversary assets and targets" are therefore also psychological tools.[84] Russian information warfare does not target solely adversarial key decision-making or armed forces personnel, but also the much wider "protest potential of the population."[85] Russia views the real battlefield to be human consciousness, perceptions, and strategic calculations.[86] And this battlefield is boundless: it blurs the boundaries between war and peace, internal and external, tactical, operational and strategic levels of operations, forms of warfare (offense and defense), and of coercion.[87]

In China, following the Chinese PLA Science of Military Strategy (2013) and the subsequent PLA reforms (2015), the role of information warfare, both in peacetime and in wartime, has simultaneously increased and has been reinforced in non-military components of CCP power, including the intelligence agencies, state-owned enterprises, and other ostensibly civilian institutions. Together, these institutions are all bound by specific strategies that have been agreed upon at the highest level to shape and form international narratives that further guarantee Chinese strategic interest in cyberspace. This is reflected in the Chinese 'Three Warfares' strategy that relies on three mutually reinforcing hybrid strategies: (1) the use of strategic psychological operations (information warfare and covert influencing), (2) overt and media manipulation (strategic communication and overt influencing), and (3) legal warfare intended to manipulate the strategies, policies, and perceptions of the target audiences abroad.[88] The control of information, including Internet content and physical infrastructure, is seen as a security warrant for the survival of the regime, much like Russia albeit with more domestically oriented goals and overt means.

Within the conventional and nuclear forms of deterrence, states are the actors being deterred. In cyber, things are a bit more complicated.

3.3. **Actors**

Within the conventional and nuclear forms of deterrence, states are the actors being deterred. In cyber, things are a bit more complicated. It covers not only a much wider range of states, and many more agencies within a single government, but also non-state actors that can act independently as cybercriminals or as state-affiliated or directed proxies – or all at once. This section introduces the main state and non-state actors and their hybrid relations.

3.3.1. **State actors**

While there has been a steady uptake of Military Cyber Commands, the game of offense remains dominated by the intelligence community and hybrid – often non-state – actors that often operate with a direct or indirect link to government. Unlike other military branches, governments have a multitude of agencies involved in cyber operations. No single agency

[84] Kiselyov, "What Kind of Warfare Should the Russian Armed Forces Be Prepared for?".

[85] Government of Russia, "The Military Doctrine of the Russian Federation," (December 25, 2014).

[86] Dimitry Adamsky, "Cross-Domain Coercion: The Current Russian Art of Strategy," Proliferation Papers 54, IFRI Security Studies Center (November, 2015).

[87] Sergei Modestov, "Strategicheskoe sderzhivanie na teatre informatsionnogo protivoborstva" [Strategic containment in the theatre of information counter-struggle], *Vestnik Akademii Voennykh Nauk, vol 26*, no. 1, (2009), cited in Adamsky, "Cross-Domain Coercion: The Current Russian Art of Strategy", op. cit.

[88] Elsa Kania, "The PLA's Latest Strategic Thinking on the Three Warfares," China Brief Volume: 16, Issue 13, Jamestown Foundation (April, 2016) https://jamestown.org/program/the-plas-latest-strategic-thinking-on-the-three-warfares/.

has the sole responsibility for cyber operations, which has also led to internal competition and various agencies conducting similar operations without apparent awareness of each other.

Intelligence agencies

For both offense and defense, there is a fundamental tension between intelligence collection and military action.

Historically, cyber operations originated from the intelligence community, often the SIGINT operators, and to this day remains the place where much of the expertise still resides, especially considering that every attack (CNA) requires espionage (CNE). For both offense and defense, there is a fundamental tension between intelligence collection and military action. The latter aims to exploit vulnerabilities to generate an effect that is likely to be seen by the target (or more widely). Intelligence agencies, on the other hand, operate covertly to extract information and maintain a long-term information position within the enemy's network.[89] This partly explains the dual-hatted function of the Commander of USCYBERCOM and the head of the NSA, which is described as "a way to navigate the normative and legal thicket of Title 10-Title 50 debates on cyber operations."[90] Within the US, the NSA Tailored Access Operations (TAO) unit has been and continues to be one of the most notable players in the field, although USCYBERCOM has arisen as the operational command in charge of offensive cyber operations. As shown in Annex II, which includes more details on the various intelligence agencies, TAO was aggressively expanded to develop a global architecture and tools that can augment its traditional passive intelligence collection into covert action and sabotage that extended to preparation of the battlefield.[91] To this end, it has shown a willingness to take high political risks in the pursuit of intelligence gains, indicative of the American path dependency in cyber operations that takes a bottom-up approach of placing the tactical before the political. In other words, there is a natural trend in favoring the technical operators in a highly complex and esoteric field as cyber and a "bottom-up culture of putting technical feasibility before political desirability, which is hardwired into the NSA and US Cyber at large."[92] This was possible through the complete absence of public discussion and relatively few constraints and congressional scrutiny on NSA's foreign intelligence activities compared to domestic ones.

Other countries, like the UK, decided to embed their wide range of cyber operations (from defensive, to intelligence and offensive) in their SIGINT agencies, namely the Government Communications Headquarters (GCHQ). Together with the Ministry of Defense, it jointly runs the National Offensive Cyber Program, which is reported to have a budget £250 million and a

[89] One telling example comes from Ash Carter, former US Secretary of Defense, who expressed his disappointment when describing that effective cyber operations against ISIS were never fully realized as "the intelligence community tended to delay or try to prevent its use, claiming cyber operations hinder intelligence collection". Ash Carter, "A Lasting Defeat: The Campaign to Destroy ISIS," Harvard Belfer Center Special Report, (October 2017), 33; Boeke and Broeders, "The Demilitarisation of Cyber Conflict", 76-77.

[90] Steven Loleski, "From cold to cyber warriors: the origins and expansion of NSA's Tailored Access Operations (TAO) to Shadow Brokers," *Intellligence and National Security, vol. 34*, no. 1 (October 2018), 123.

[91] Traditionally, the US' strategies and budgets appear to favor offensive over defensive. Programs that appeared as purely defensive to the outside, actually heavily leaned towards offensive measures, especially the capabilities of the NSA. See for example the 2008 Comprehensive National Cybersecurity Initiative (CNCI) that initially appeared to be purely defensive and made reference that would make outsiders consider it to be offensive. From the Snowden leaks, however, it became clear that served as the direct justification for an annual $650 million USD expansion of NSA operations direct at securing presence on endpoints. TAO's endpoint activities largely rest on obtaining remote access via covert implants and infrastructure throughout the world. To illustrate, in 2004, NSA was managing about 100–150 implants worldwide to 21,252 by 2008 and was projected to control 85,000 by the end of 2013. Klimburg, *The Darkening Web,* 153; Ryan Gallagher and Glenn Greenwald, "How the NSA Plans to Infect 'Millions' of Computers with Malware", The Intercept, (March 12, 2014); Barton Gellman and Ellen Nakashima, "U.S Spy Agencies Mounted 231 Offensive Cyber-Operations in 2011, Documents Show," Washington Post, (August 30, 2013).

[92] Klimburg, *The Darkening Web,* 149-150.

staff of 2,000 in 2018.[93] It also houses the previously described JTRIG and a Human Science Operations Cell that engages in online covert action focusing on online human intelligence, strategic influence, disruption and computer network attacks with the aim of understanding and manipulating the wider online discourse.

But in no other country is the role of the intelligence community in carrying out cyber operations so prominent as in Russia. Most notably, this includes its foreign intelligence agencies, the GRU (military), and SVR (civilian). The GRU is considered to have the best technological capabilities among all Russian special services and is well-known for carrying out "Russia's most brazen and damaging cyberattacks", including the 2015 attack against Ukraine's electrical infrastructure, the 2016 US presidential election and the 2017 French presidential election, the 2017 NotPetya ransomware attack, and the 2018 hacking attempt against the OPCW amongst many others.[94] It is difficult to compare the GRU to agencies in other nations as it not only houses usual forms of intelligence and cyber operations (unit 26165 and 74455), but also disinformation units (Unit 54777) and the special-operations forces used for peacetime covert action as well as wartime sabotage missions (the infamous *Spetsnaz* troops).[95] The SVR is the primary civilian foreign intelligence agency that mostly targets government networks, think tanks, and information technology companies, and was found to be behind the SolarWinds operation.[96] APT29 – also known as CozyBear or the Dukes – has been linked to the SVR. While some of the Russian intelligence agencies are more focused on building their internal capabilities than others, the most defining organizing characteristic is the hybrid nature in which they operate, closely with outside contractors, cybercriminals, and hacktivists.

The Western states, the 'militarization' of cyber signified a 'de-spookification' effort of cyber operations in which it is integrated within the larger concept of military operations.

Military Cyber Commands

Following the lead of the US in establishing their Cyber Command in 2010, a growing number of states have established military cyber commands to conduct strategic cyber operations and integrate them within their national arsenal. The most apparent reason for this trend is to ensure security in a world that is increasingly digitalized and therefore vulnerable to exploitation (in particular by other nation-states). Unlike other military systems, however, cyber operations were largely carried out by the intelligence community and treated as dark secrets from the intelligence world of SIGINT, lacking much-needed transparency.[97] Indeed, for most of the Western states, the 'militarization' of cyber signified a 'de-spookification' effort of cyber operations in which it is integrated within the larger concept of military operations. Another explanation for this trend is the socialization effect that early movers, such as the US Cyber Command, had on its like-minded partners.[98] Before these commands were set up, most military branches primarily focused on cyber and electromagnetic activities (CEMA) and to a certain extent information warfare capabilities.

[93] The Telegraph, "Britain steps up cyber offensive with new £250m unit to take on Russia and terrorists;" Telegraph.co.uk, (September 21, 2018).

[94] Congressional Research Service, "Russian Cyber Units," CRS Reports, (January 4, 2021).

[95] From its mission-specific directorates, the Sixth Directorate, in charge of electronic/signals intelligence, hosts the notorious Unit 26165 and Unit 74455. Unit 26165 was established as the 85th Main Special Service Center responsible for military intelligence cryptography during the Cold War. Unit 74455, on the other hand, appears to be a much more recently-established unit to help expand GRU cyber capabilities. Also known as the Main Center for Special Technologies or for outsiders as *Sandworm*, the Unit was indicted by the DoJ for a number of cyber operations. Finally, there is Unit 54777, known as the 72nd Special Service Center, that is responsible for psychological operations, more recently linked to online disinformation campaigns targeting the Covid-19 pandemic. They provide support to the other cyber units and also operate on the tactical level by harmonizing electronic warfare and information warfare operations.

[96] United Kingdom National Cyber Security Centre, "UK and US call out Russia for SolarWinds compromise, (April 2021).

[97] James A. Lewis, "The Rationale for Offensive Cyber Capabilities," CSIS, (June 8, 2016).

[98] Max Smeets, "Cyber Arms Control: Incentives and Challenges," The Hague Program for Cyber Norms, (2021).

In the US, USCYBERCOM is the leading military organization with a cyber mission force as its operational arm. The force consists of Combat Mission Teams (CMTs), which operate with the combatant commands to support their missions, and Combat Protection Teams (CPTs), which defend DoD Networks, and the Cyber National Mission Force (NMF). All three continue the trend of combining both attack and defense roles. As the only body directly under the direction of USCYBERCOM, "the NMF is a unit co-located with the NSA that has previously played a crucial role in operational cyber, but which seems to have been heavily pared back under the current strategy."[99] NMF's National Mission Teams appear to be "the true carriers of strategic cyber operations, able to strike back not only at the adversary's cyber or conventional forces but also at its government and civilian infrastructure".[100] What is less clear, however, is that there is no public indication of what these teams could do, practically or legally speaking. In fact, the introduction of "Offensive Cyber Effects Operations"(OCEO) expanded the level of ambiguity of what these forces are intended to do.

Within China, the People's Liberation Army (PLA) fulfills the leading role when it comes to offensive cyber operations. Traditionally, PLA strategic thinking focused on 'information dominance' through operations, targeting the adversary's command and control systems and using integrated information and firepower assaults.[101] To this end, the PLA mostly concentrates on information operations that include cyberwarfare, electronic warfare, and psychological warfare. In an effort to synchronize these operations, the Strategic Support Force (SSF) was established in 2015-16 as part of a massive PLA reform.[102] It signified an important push towards the militarization of previously intelligence-driven PLA capacities. SSF integrated the former stovepiped Chinese military cyber, electronic, and information warfare capabilities. This was done by consolidating the technical reconnaissance bureaus from the 3PLA and the electronic and offensive cyber capabilities from the 4PLA with the aim of developing more significant cyber fires.[103]

In cyberspace, the monopoly of violence by the state is challenged by the dominant role of non-state actors in various shapes and forms, as well as their unclear relationships with governments.

3.3.2. **Non-state Proxies**

A government's intelligence and military branch operate under different legal regimes, but at the very least they operate as recognized state actors that are supposed to work within the bounds of the law. In cyberspace, the monopoly of violence by the state is challenged by the dominant role of non-state actors in various shapes and forms (attacker, victim, medium, or carrier of attacks), as well as their unclear relationships with governments. When Estonia, in 2007, was hit by what has sometimes been called the first strategic cyberattack in history – a DDoS attack paralyzing its government, media agencies, and financial institutions - it marked a watershed moment in the use of state-sanctioned cyberattacks to advance foreign policy

[99] Klimburg, *The Darkening Web,* 154.

[100] Ibid.

[101] Elsa B. Kania, "China's quest for political control and military supremacy in the cyber domain" CNAS (March 16, 2018).

[102] Many of the SSF forces are organized in "bases", a form of corps leader grade unit that is distinct to the PLA, of which 311 Base is known as the Chinese 'Three Warfares Base' from the General Political Department that is publicly known for its focus is on psychological warfare, although very little is known about its exact position within the SSF organizational structure. Next to the SSF, the PLA reforms also impacted the command structure. The CMC's new Joint Staff Department (JSD), formerly under the General Staff Directorate, supervises joint operations and oversees various components of military command, including operations, intelligence, cyber, and electronic warfare, communications and battlefield environment support, albeit unclear what the exact division of responsibilities between JSD and SSF.

[103] The Network Systems Department even maintains the former 3PLA headquarters, location, and internal bureau-centric structure. In at least one instance, the NSD has been referred to as the "SSF Third Department".

goals. It also introduced a model for conflict in cyberspace fought by proxy to retain some degree of plausible deniability – even when there is an overall consensus saying otherwise.

Non-state actors involved in cyber operations take on various forms that can have a formal, informal, or no relationship with a government. There have been numerous efforts to structure these relationships. Tim Maurer identifies three main relationships: *delegation, orchestration, or sanctioning*.[104] The relationship between the government and the proxy, and the latter's use, depends on a range of factors, including the domestic landscape (public-private cooperation, crime levels, etc.); the government agencies' preexisting relations with proxies; and their definition of *cybersecurity* or *information security*, where China and Russia put more emphasis on the content of data as a potential threat to domestic stability.

While the blurring between both groups is predominantly a Russian characteristic (which maintains close and fluid relations with criminal enterprises), any country will have some degree of the so-called revolving door in which parts of its cybersecurity workforce oscillates between government agencies and non-criminal private entities.

Delegation presumes a state's effective control over the proxy to which it hands over certain cyber operations. It is mostly used to describe the US government's relation to cybersecurity and intelligence companies and contractors. Formalized in contracts, it is the most formally framed, meaning they are relatively constrained.[105] It should be noted that many of the cybersecurity firms operating from the US – which a lot of these companies do – are for the most part not considered to be government proxy agencies, but independent private entities that still contribute to the overall deterrence posture. They provide a talent base for the intelligence and military agencies that are increasingly contracted in from industry (instead of tasks being outsourced to them). They also attribute adversarial transgressions as well as provide useful technical intelligence and evidence that can be used to inform attributions of the US or allies. While the blurring between both groups is predominantly a Russian characteristic (which maintains close and fluid relations with criminal enterprises), any country will have some degree of the so-called revolving door in which parts of its cybersecurity workforce oscillates between government agencies and non-criminal private entities. Finally, a re-occurring development that deals with the role of private actors in cyber offense is the so-called 'hack back' proposal within the US legislature, wherein offensive cyber operations by non-state actors in the name of self-defense would be allowed. While 'active cyber defense' by the private sector is unlawful in most states, including the US, it may be reconsidered as a lawful tool.[106] Many policy and legal questions remain, such as determining the level of confidence needed for attributing an attack before taking proportional actions, as well as defining what the latter would look like.[107]

Orchestration means a state actively backing a non-state actor, often with financial or logistical means. The Iranian government, for example, has provided financial support to students for carrying out cyber operations against the US, while the non-state Syrian Electronic Army (SEA), often described as the Syrian government's loosely governed elite cyber militia, was behind hacks of Western media outlets, human rights organizations, communications platforms, and US military websites. Interestingly, after the SEA disappeared in 2016, it resurfaced a year later in a different form, moving its focus from covert intelligence operations to a public relations extension of the government that seeks to spread disinformation and shape media

[104] Tim Maurer, *Cyber Mercenaries. The State, Hackers and Power* (Cambridge: Cambridge University Press, 2018).

[105] Ibid.

[106] In 2017, the Active Cyber Defense Certainty Act was introduced in the US House of Representatives but failed to gain traction. A similar bill now resurfaced in a bipartisan proposal. Tom Graves, "Active Cyber Defense Certainty Act," Pub. L. No. H.R. 3270 (2019); US Senate media, "117th United States Congress 1st Session".

[107] Global Commission on the Stability of Cyberspace, "Additional Note to the Norm against Offensive Cyber Operations by Non-State Actors," (November 2018).

narratives.[108] Russia is described as a country that uses both orchestration and sanctioning in its relations to proxies.

Sanctioning implies passive support or inaction by turning a blind eye to the proxy. This is arguably the largest category. In contrast to the much tighter restrictions and direction that the Chinese government places on its non-governmental actors, Moscow often stops short of directing non-state actors and allows criminal groups to carve out their path as long as they generally work towards Putin's goals. [109] The partnership between Russian cybercriminals and the intelligence community is one of convenience – cybercriminals offer resources (in particular recruitment) and infrastructure that is also useful for government cyber operations.[110] It is also a politically advantageous partnership that offers the Russian government a degree of plausible deniability as it hides behind criminal actors.[111] The availability of Russian proxies has been mobilized quickly for patriotic purposes, such as in support of Moscow's operations against Estonia (2007), Georgia (2008), and Ukraine (2014). An added bonus is that these criminals offer 'noise' under which the more skilled government hackers can move undetected. The defining factor of the Russian 'information counter-struggle' is that it is executed by a 'Whole of Nation'-approach, much like the Soviet-era notion of 'total defense' which not only encompassed government entities but all national resources. This corresponds to the description by Russia expert Mark Galeotti of how the Kremlin carries out this approach by outsourcing to volunteers, organized-crime groups, businesses, government-organized non-governmental organizations, the media, and other actors in the deployment of various active measures.[112]

The intelligence community is one of convenience – cybercriminals offer resources (in particular recruitment) and infrastructure that is also useful for government cyber operations.

Finally, China is described as having a state-proxy relationship that moved from sanctioning to orchestration, and eventually delegation. The Chinese government's increasing control over proxy actors, exercised via traditional militia groups or patriotic hackers, coincided with an incremental hardening of Chinese Internet governance and control. IP theft campaigns were mainly carried out by non-state forces and were likely a useful way to keep these forces busy and their attention focused on outsiders rather than on domestic – in particular government – targets. Government actors were not only hiding in the noise created by the non-state actors (at least until the Xi-Obama agreement in 2015 condemning cyber-enabled economic espionage), but actively encouraging civilian attacks as well.[113] Clearly, Chinese authorities exercise some degree of control over at least some of the non-governmental hacking groups, albeit not always clear to what extent the activity was actually directed, rather than simply encouraged or tolerated. Similar to Moscow, Beijing brings outside hackers into the government fold and is known for its fusion between military and civilian entities.

[108] It has been reported that "offensive cyber operations continue, but overall the SEA appears less technically sophisticated and more concerned with shaping the media narrative, disinformation and restraining the public's online behavior. The new SEA includes a media office and regional offices in various Syrian governorates." Abdulrahman Al-Masri and Anwar Abas, "The new face of the Syrian Electronic Army," Opencanada.org (May 17, 2018).

[109] More specifically, a distinction is made between three types of associations between the intelligence services and criminal groups: direct links (e.g. the case of Dmitry Dokuchaev – a former cybercriminal who was recruited by the FSB), indirect affiliations (e.g. GameOver Zeus botnet) and tacit agreement (activity without a clear link but allowed by the Kremlin, which turns a blind eye to it). The report found that it is very unlikely that these associations and activities will come to an end, although they may adapt to provide greater plausible deniability through fewer overt and direct links between the spooks and criminals. Recorded Future Insikt Group, Cyber Threat Analysis Russia," (September 2019).

[110] Klimburg, *The Darkening Web.*

[111] Andrei Soldatov and Irina Borogan, "The Red Web: The struggle between Russia's digital dictators and the new online revolutionaries," *Journal of Strategic Security, 8*(4): 122 (2015).

[112] Mark Galeotti, "Putin's hydra: Inside Russia's intelligence services", European Council on Foreign Relations, (May 11, 2016).

[113] Klimburg, *The Darkening Web*, 288.

3.4. **Thresholds**

Thresholds in cyberspace are difficult to define as cyber operations are often scalable or reversible in ways that conventional operations are not. A legal approach is taken to describe the waves in which thresholds were established in cyberspace. Initially, Western efforts were directed at the physically destructive effects of cyber operations. States faced a lack of consensus as to what constitutes a 'cyber attack'. In fact, some disagreement remains between the so-called East and West. One solution was to look at the long-standing international norms defining an armed attack. The Tallinn Manual, a non-binding but authoritative academic manual, consequently defined a cyber attack as "an equivalent to an armed attack, with significant casualties and/or economic loss, or weakening of national security."[114] In this effects-based approach, cyber operations can thus reach the threshold of an armed attack by causing a loss of life and significant economic harm, which has been reaffirmed by a growing number of like-minded states, including the Netherlands and France.[115] This means that states are therefore entitled to the use of force as a self-defense measure as defined in Article 51 of the UN Charter, which may take on any other form as long as it abides by certain conditions. Similarly, a state is allowed to take countermeasures when it is a victim of use of force from another state.[116] The UN Charter does not define an armed attack or use of force, but a definition has crystallized over the years through case law that defines an armed attack as the most serious form of the use of force on the axes of scale and effect, and usually involve some form of physical injury or damage.[117]

Even these relatively well-established thresholds are not always clear to determine in the cyber context. Loss of life is the highest threshold that, as Libicki puts it, in terms of clarity, has the advantage of being unambiguous.[118] Considering the chain of events that would lead up to it (in peacetime), he believes it is "likely to come because some accident was made more likely or because some warning and control system was knocked offline. Given the indirect chain of events cited here, justifying retaliation based on such an event would hardly be simple."[119] Turning to economic criteria – e.g. attacks that cost more than 1 million euros – he adds that they are tractable and offer some reasonable proportionality, but again are hard to define. It would also establish a double burden for the retaliator who "not only needs to establish causality between an attack and the subsequent damage but also, unless the threshold was low or the damage clearly high, make a convincing case that the damage exceeded the threshold."[120]

[114] Michael N. Schmitt, *Tallinn Manual on the International Law Applicable to Cyber Warfare* (Cambridge: Cambridge University Press, 2013).

[115] Government of The Netherlands, "Appendix: International Law in Cyberspace", (26 September, 2019); Ministére des Armées, "International Law Applies to Operations in Cyberspace", (24 September, 2019).

[116] The prohibition on the use of force, stipulated in article 2(4) of the UN Charter, was also left undefined but is measured across the same axes of scale and effect. Article 2(4) of the UN Charter states that "all Members shall refrain in their international relations from the threat or use of force against the territorial integrity or political independence of any state, or in any other manner inconsistent with the Purposes of the United Nations." United Nations, "Charter of the United Nations," (10 August 10, 2015). International Law Commission, "Draft articles on Responsibility of States for Internationally Wrongful Acts, with commentaries," *United Nations,* Part III, chapter II, 2001.

[117] See the Nicaraguan case. This mostly has to do with the scale and effects, although international law is ambiguous on the precise thresholds for this. Duncan Hollis, "New Tools, New Rules: International Law and Information Operations", in The Message of War: information, Influence and Perception in Armed Conflict, G. David and T. McKeldin, eds., 2008, 63.

[118] Martin C. Libicki, "Cyberdeterrence and Cyberwar," RAND.org (2009), 67.

[119] Ibid.

[120] Ibid. 68

At the same time, the Tallinn Manual stipulated that "non-violent operations, such as psychological cyber operations or cyber espionage, do not qualify as attacks."[121] This is not supported by Russian and Chinese interpretations of the use of force which include psychological and media warfare.[122] Their perceptions of information as a weapon consider bad *content* as critical or dissenting of the regime and thereby an attack against the state. Furthermore, at least China has been concerned that a use of force could be expanded to include wider array of cyber operations that could invoke countermeasures against economic espionage on the basis of its severe economic damage. Related to this is an inherent issue of the effects-based approach. It leaves no "logical basis" to exclude that which has traditionally always been excluded from the prohibition on the use of force: economic coercion.[123]

Russian and Chinese cyber operations exploiting the gray zone have generally sought to test the response thresholds of their opponents.

Russian and Chinese cyber operations exploiting the gray zone have generally sought to test the response thresholds of their opponents. Nonetheless, they steer clear of causing physical harm, at least in cyberspace, and thereby from tripping over the armed attack and use of force threshold that would warrant self-defense or countermeasures. Given the nature of cyber operations and the dependency on digital technologies for many of its basic functions, operations are likely to be more disruptive than destructive. Fisherkeller and Harknett, when describing the specification of thresholds and the shortcomings of deterrence, stated: "An alternative, non-territorial-based cyberspace threshold has been discussed that, if crossed, would justify responses of armed attack, but that threshold is based on cyber OAAs that cause the damage equivalent to the use of force. Such damage is of a very different nature and not representative of the significant damage being caused by on-going espionage, sabotage, and subversion cyber OAAs and, consequently, does not shape this consequential adversarial behavior occurring regularly below the use of force threshold."[124]

Moving below these thresholds, the next legal barriers would be the principles of nonintervention and sovereignty. The latter offers a good starting point but yields little relief by itself given the ongoing debate among like-minded states as to whether sovereignty itself is an enforceable rule or merely a principle of international law.[125] The principle for nonintervention in the internal affairs of other states is also well-established within customary international law. It allows states to safeguard their sovereignty and independence, and its application to cyberspace has been established and reinforced by many states.[126] Like the use-of-force prohibition, the nonintervention rule is considered to be of limited scope. Fundamentally, it

[121] Schmitt, *Tallinn Manual on the International Law Applicable to Cyber Warfare*

[122] Taylor Cruz; Paulo, Simoes, "*EECWS 2019 18th European Conference on Cyber Warfare and Security*", Academic Conderences and Publishing Limited, (4 July, 2019),

[123] L.J.M. Boer, "Restating the Law "As It Is": On the Tallinn Manual and the Use of Force in Cyberspace," *Amsterdam Law Forum vol. 5* no. 3, (2013), 10.

[124] Michael P. Fisherkeller and Richard J. Harknett, "Deterrence is Not a Credible Strategy for Cyberspace," *Foreign Policy Research Institute* (2017), 386. doi: 10.1016/j.orbis.2017.05.003

[125] Within the cyber context, there is an ongoing debate as to whether sovereignty itself is an enforceable rule of international law or merely a principle of international law. France is among the former group and holds that "any unauthorized penetration by a state into French systems or any production of effects on French territory via a digital vector may constitute, at the least, a breach of sovereignty". Austria, Germany, the Netherlands and the Czech Republic also agree with the sovereignty-as-a-rule interpretation, albeit with varying degrees as to what kind of activity would automatically constitute a violation of sovereignty. By contrast, the US, like the UK, holds the view that sovereignty is merely a principle of international law and does not create autonomous and separate legal obligations, but is protected by other established rules of international law, such as the prohibition of the use of force or the principle of non-intervention. Przemyslaw Roguski, "The Importance of New Statements on Sovereignty in Cyberspace by Austria, the Czech Republic and United States", Just Security (11 May 2020); Ministry of Defense France, "International Law Applied to Operations in Cyberspace." (2019)

[126] The International Court of Justice (ICJ) has described the principle of non-intervention as "a corollary of every state's right to sovereignty, territorial integrity and political independence," and of the right, as a matter of sovereign equality, of every state to conduct its affairs without outside interference. International Court of Justice, "Case Concerning Military and Paramilitary Activities in and Against Nicaragua", (1986).

prohibits the use of *coercive* measures, the most coercive being the use of force, to overcome the free will of a targeted state with respect to matters that fall within that state's core, independent sovereign prerogatives.[127] "Unfortunately, the concepts of coercion and 'domaine réservé'—the bundle of sovereign rights protected by the rule—are ill defined".[128] Such ambiguities can be cleared up by states disclosing their official views and interpretations. Thus far, only a handful of states have done so. The most concrete statements that go beyond a general acknowledgment that the parameters of the principle 'have not yet fully crystallized in international law' is meddling in electoral processes.[129]

Accompanying and expanding on the existing legal understandings and restrictions, states embarked on a path of norm development in an effort to proscribe redlines. Within the cyber context, this process was initiated through the UN Group of Governmental Experts on Developments in the field of information and telecommunications in the context of international security (UN GGE). Notably, in 2015, the UN GGE presented 11 norms, which included commitments such as requiring states to not knowingly conduct or support wrongful acts in cyberspace, including actions that intentionally damage or impairs critical infrastructure or targets computer emergency response teams.[130] These norms are broadly adopted by all members of the UN General Assembly.

Despite these normative and legal thresholds, threats still exist, and cyber operations continue to take place. Adversarial operations try to avoid detection, bypass existing norms, laws, and response thresholds to undermine the basis of decisive response. They also intentionally inch closer to test where the barriers for violations lie. Two well-known examples come to mind. First, the BlackEnergy operation from December 2015, attributed to Russian APT group Sandworm, signifies the first-of-its-kind attributed attack paralyzing Ukrainian electricity grids, leaving more than 230,000 residents in the dark.[131] It would also be the first breach of one of the most important UN GGE norms that prohibits operations against critical infrastructure adopted and agreed by the UN General Assembly earlier that year. These norms govern peacetime operations, whereas this operation was carried out as part of an ongoing hybrid conflict in Ukraine. Second, the cyber operations of APT-28 - aka *Fancy Bear* - between 2016 and 2018, operating as part of Russia's GRU, targeted US and French political parties and election processes.[132] This operation could again be considered a violation of the

[127] Interventions against the sovereignty and the principle of non-intervention require an element of coercion. This concept can be defined broadly or narrowly, with great consequences for the analysis of the case. Unfortunately, international law says very little about the theory of coercion. A complete analysis of what constitutes coercion within this context of international law is too expansive for this study. For more information about this, see Jens David Ohli", "Did Russian Cyber Interference in the 2016 Election Violate International La"?," *95 Texas Law Review* 1579 (2017); Duncan B. Hollis, "The Influence of War; The War for Influence." SSRN Scholarly Paper, Social Science Research Network, (3 April, 2018).

[128] Gary Corn, "Coronavirus Disinformation and the Need for States to Shore Up International Law", *Lawfare* (April 2, 2020).

[129] The Netherlands referenced to the principle of non-intervention when it called out Russian disinformation campaigns during the COVID-19 pandemic. UNODA. "The Kingdom of the Netherlands' response to the pre-draft report of the OEWG" (April 2020); Corn, "Coronavirus Disinformation and the Need for States to Shore Up International Law".

[130] Henry Rõigas and Tomáš Minárik, "2015 UN GGE Report: Major Players Recommending Norms of Behaviour, Highlighting Aspects of International Law," CCDOE, (2015).

[131] Kim Zetter, "Inside the Cunning, Unprecedented Hack of Ukraine's Power Grid," Wired.com, (March 3, 2016).

[132] Hacking of electoral infrastructure and parties in the US presidential elections from March 2016, primarily directed at the Democratic National Committee (DNC) Clinton's campaign, and subsequently the French elections in 2017, which targeted the Macron campaign. The attack methods centered on spear phishing campaigns to capture user credentials in order to access and subsequently leak confidential documents; overtly monitor the computer activity of dozens of employees; and implant hundreds of malicious files to steal passwords and maintain access to the networks. The leaked documents were altered with fabricated information, amplified through Russian-aligned media outlets, such as RT and Sputnik, internet trolls, and co-opted sympathetic groups, like Wikileaks.

same UN GGE norm but the norm does not specify what constitutes critical infrastructure, and at the time the US did not label its electoral infrastructure as critical, although this was later added.

Thresholds are not just defined legally or normatively, but also by actions and responses from states that are informed according to their essential national interests, the latter being explicitly articulated or not. For Russia, an articulated national interest was to slow down NATO's Eastern expansion. In the case of the US, national interests are not always clearly defined, allowing it to retain some degree of strategic ambiguity. One of the strongest US responses to China thus far was, for example, triggered by economic espionage. Indictment and the threat of sanctions created sufficient leverage for bilateral negotiations to mitigate reciprocal escalation through the establishment of a Memorandum of Understanding prohibiting cyber-enabled IP-theft for economic gains, generally known as the Xi-Obama Agreement.[133] In doing so, Washington may have hoped to make clear the distinction between 'acceptable' and 'unacceptable' espionage. After all, the US limited the norm to cyber-enabled IP theft for *economic benefits*. The underlying hope was to get China to accept a distinction between legitimate traditional espionage for political-military ends and illegal espionage for commercial ends.[134] The Agreement had a very clear impact at the start,[135] with heavy reduction of Chinese-attributed espionage, a decrease however which proved temporary as tensions increased markedly under President Trump from 2017 onwards. Why the original decline happened is not completely clear, even for individuals directly involved in the agreement. The implied threat of sanctions may have been a sufficient threat on its own. Alternatively, China may have rationalized its actions by bringing relatively 'noisy' bulk espionage activity under a more concise and manageable structure, and with a higher level of proficiency. The Chinese leadership may have also simply used the crisis as an opportunity to bring parts of the PLA to heel.

Thresholds are not just defined legally or normatively, but also by actions and responses from states that are informed according to their essential national interests, the latter being explicitly articulated or not.

"While some ambiguity can be helpful in deterrence", Michael Daniel warns that "too much ambiguity reduces its utility. Daniel argues that deterrence already works to keep states from using their offensive cyber capabilities to cause widespread, frequent disruption and even physical destruction against critical infrastructure outside of armed conflict "as nations such as the US, the United Kingdom, Israel, China, Russia, and Iran used offensive cyber capabilities to this end already, they could choose to do so on a regular basis, but they do not do so".[136] Below this threshold, deterrence is still possible and can work – at least in some areas. In what he describes as expanded deterrence, Daniel explores how cyber deterrence can be expanded to cover (parts of) the gap in between the part where cyber deterrence currently works (against critical infrastructure) and where it will not work (cyber espionage). He encourages the US and its allies to determine what specific behavior they deem unacceptable: "To date, the US and its allies have not clearly tied deterrence efforts to behavioral benchmarks. Such benchmarks would not constitute redlines (as in, if you do x, we will do y), but rather an articulation of what malicious cyber activities the US and its allies seek to deter beyond what is already deterred." [137] Therefore, reducing ambiguity in the actions they want to deter "does not

[133] White House, "FACT SHEET: President Xi Jinping's State Visit to the United States," *Obama White House, (*September 25, 2015).

[134] Adam Segal; Samantha Hoffman; Fergus Hanson; Tom Uren, "Hacking for Ca$h", ASPI, (2018).

[135] Louk Faesen et al, "From Blurred Lines to Red Lines. How Countermeasures and Norms Shape Hybrid Conflict," *The Hague Centre for Strategic Studies* (September 2020); FireEye iSight Intelligence, "Redline Drawn: China Recalculates its Use of Cyber Espionage," *Mandiant* (June 2016).

[136] Michael Daniel, "Closing the Gap: Expanding Cyber Deterrence," *Cyberstability Paper Series. New Conditions and Constellations in Cyber* (July 2021), 3.

[137] Ibid, 7.

require committing to a specific action in response to such behavior, but effective deterrence does require a consistent overall response to such activities."[138]

3.5. **Main Takeaways**

For SMPs to design a deterrence posture in cyberspace that includes deterrence by punishment, a proper understanding of the actions, actors, and escalation thresholds is required. To that purpose, this report introduces a typology for cyber operations including passive and active measures. Passive measures are defensive in nature and take place in friendly or neutral networks, while active measures are deployed in the enemy's networks. Active measures are categorized as: (1) Cyber Surveillance and Reconnaissance (ISR), or covert intelligence operations; (2) offensive counter cyber operations, aimed at impairing a target enemy's malicious activity; (3) cyber support operations, deployed in support of other civilian and military operations; and (4) cyber attack operations. While intelligence operations are generally accepted by states, in cyberspace they could potentially lead to inadvertent escalation since for the targeted state distinguishing between a cyber ISR and *preparation-for-battlefield* is very difficult.

A distinct category of cyber operations is information warfare, pursued mainly by non-Western states for content-control purposes to safeguard their political regimes and influence the adversary's perception of reality. Although some like-minded countries are increasingly recognizing the importance of influence and information operations, it is not yet considered to be a war-winning strategy in and of itself. The cyber order of battle remains primarily focused on kinetic-equivalent effects which impact the way information warfare capabilities are being used as a supportive tool. Russia, and China to a certain extent, uses a different paradigm with information warfare and psychological effects as to the desired end goal to which physical and kinetic operations can contribute.

Traditionally, both conventional and nuclear deterrence have focused on states as the key actors being deterred and doing the deterring. Cyber deterrence is significantly more complicated because of the multitude of actors, both state and non-state. First, deterrence may be targeted at fully independent non-state actors. Second, the actual affiliation of actors can be multiple all at once (government, proxy, and rogue actor). Third, states are not monolithic entities and many different departments can engage in cyber operations, often leading to a cacophony of action, not only from varying mandates within governmental but also due to the activities of proxies and other state-affiliated organizations.

Likewise, they need to be influenced through different means and ends. Still, deterrence can be done by and against a larger a number of actors. Hence, a reasonable level of attribution becomes a crucial component of a successful deterrence strategy. Finally, the thresholds in cyberspace are not as clearly established as in other areas. While thresholds can be clarified by publishing national doctrines or statements, by implementing norms of responsible behavior, and by acknowledging a country's actions and responses, only a few countries have done so. Although it is true that ambiguity within deterrence can be useful to maintain freedom of action, too much is detrimental to deterrence's utility. Thus far, states have not been very successful in establishing thresholds, but also at linking specific retaliatory measures to those thresholds.

[138] Ibid.

Related to the risk of unintended second- and third-order effects described in the previous chapter, when it comes to cyber attack operations, two components need to be considered: the level of preparation and how to deal with ambiguity and unintended effects. The careful preparation of a cyberattack will significantly enhance its success, as the level of preparation is directly proportionate to the protection level of the target and the distinction (limitation) of the preferred effects.

Dealing with ambiguity is not only a feature of cyber attacks, but of cyber conflict writ large.

But even the best preparation will only go some way in addressing the second distinguishing feature: managing ambiguity and inadvertent and unintended effects. This challenge starts with limiting the direct effects of an attack, signaling attribution and ownership of an attack in a controllable manner, considering implications for other policy goals, on friends and allies, and even the Internet (Internet governance) itself, and finally, the impact on international law and the evolving rules-based world order. Dealing with ambiguity is not only a feature of cyber attacks, but of cyber conflict writ large. Therefore, not engaging in cyber operations (in particular in response) impacts each of these elements. Nonaction is very much action in cyberconflict.

4. Retaliation Means and minimum cyber deterrence

4.1. Introduction by Herbert Lin

This section focuses on deterrence of high-end cyberattacks—cyberattacks that have large-scale strategic effects on a short time scale, where the "short time scale" qualifier excludes multiple cyberattacks with long-term cumulative effects, such as cyberattacks that are focused on the theft of intellectual property. A canonical example of a high-end cyberattack would be a large-scale offensive operation directed at incapacitating the electric grid of a nation for extended periods of time.

A number of different responses are possible. Over the past decade, responses to adversary cyber operations have generally involved diplomatic, economic, and law enforcement actions. Diplomatic actions include demarches, complaints issues through bilateral channels, public attribution and naming and shaming, and diplomatic expulsions. Economic actions include financial sanctions (e.g., asset freezes), trade embargoes (flight and shipping bans, limitations of export and access to markets), foreign assistance reductions and cut-offs, arms embargoes (prohibition of weapon and dual-use exports), and travel restrictions (visa bans). Law enforcement responses generally involve domestic indictments. When those indicted are abroad as they usually are, they are beyond the reach of domestic law enforcement, though extradition is sometimes a possibility. Extradition is a substantial threat that sharply limits the ability of an indicted individual to travel outside his or her own country.

When high-end cyberattacks are at issue, diplomatic, economic, and law enforcement responses are likely to be regarded as ineffective. On the other hand, kinetic military responses may not be feasible without risking significant escalation or against an adversary that is much stronger militarily. For such reasons, cyber responses would often be preferred by national leaders.

Cyber responses can be directed at counterforce, countervalue, and counterpolitical targets. Counterforce targets are usually those associated with military assets and successful counterforce cyber operations degrade adversary military capabilities to some degree for some length of time. Countervalue targets are strategically significant assets with broad economic or societal value but are not associated with the military; successful countervalue cyber operations degrade important civilian functions to some degree for some length of time. Counterpolitical targets are usually nonmilitary as well but are characterized by being of high value to specific politically important individuals or organizations; successful counterpolitical cyber operations cause targeted damage to those individuals or organizations and may not be widely known throughout society.

Counterpolitical targeting has often been the responsibility of the operational divisions of intelligence services. In the United States, such activity would generally be associated with covert action, which is defined as an activity to "influence political, economic, or military conditions abroad" where it is "intended that the role of the United States Government will not be apparent or acknowledged publicly." In the past, for example, US covert action has been used to influence the outcome of foreign elections.[139] In principle, covert action could be used to target the wealth of foreign leaders or their close associates. Because they can be conducted with plausible deniability and calibrated in the scope and scale of their effects, cyberattacks are ideally suited to be instruments of covert action.[140] In the lexicon of this report, covert actions employing cyberattacks or offensive cyber operations result in "special" cyber effects.

Because they can be conducted with plausible deniability and calibrated in the scope and scale of their effects, cyberattacks are ideally suited to be instruments of covert action.

By contrast, counterforce and countervalue targeting would be the responsibility of the military. If the aggressor is a peer nation, the victim would be able to consider response options that were comparable in scale to the attack perpetrated on it. But smaller nations are without the resources to respond in kind at a comparable scale, and here deterrence may rest on the ability to persuade a more powerful adversary that, should it attack, it would suffer a level of damage that would exceed the value represented by such an attack—deterrence of the strong by the weak—an approach based on the deterrent value of being able to tear off an arm from the would-be attacker.[141]

What considerations must be taken into account for a nation less powerful in cyberspace to be able to execute a special or strategic retaliatory strike in cyberspace against an aggressor with superior power? A first consideration is the strength and robustness of the aggressor's defenses. In particular, counterforce targets are likely to have better cyber defenses than countervalue targets. In principle (though sometimes not in practice), the retaliator can expect to encounter both military discipline in personnel to carry out appropriate security procedures and military-grade cybersecurity technology, both of which are less available for most targets in civilian sectors.

The weaker nation is also likely to concentrate its resources for retaliatory cyber operations on one or a few individual civilian sectors to achieve significant effects rather than distributing them over a large number of sectors. In other words, crippling one critical sector will have a greater effect on an adversary than modest damage across many sectors.

Concentration on one or a few sectors also conserves intelligence resources. Offensive cyber operations depend on a very high level of intelligence support because of the strong coupling between targets and cyber weapons. Because even a very small change in the cyber posture of a target (e.g., the installation of a vendor's software patch) can negate a penetration attempt, those conducting an offensive cyber operation must have very current knowledge of the target's composition and configuration. In addition and in general, access to a cyber

[139] David Shimer, "When the CIA Interferes in Foreign Elections," *Foreign Affairs,* (June 21, 2020)

[140] National Research Council, *Technology, Policy, Law, and Ethics Regarding U.S. Acquisition and Use of Cyberattack Capabilities* (Washington D.C: The National Academies Press, 2009) https://doi.org/10.17226/12651.

[141] In the nuclear domain during the Cold War, this was the approach taken by Britain and France in deterring the Soviet Union from nuclear attack. See, for example, Josef Joffe, "The Cost of Abandoning Europe," *The National Interest*, no.3 (1986), 30-42.

target is established before that access is exploited to conduct a cyberattack against it. These accesses must be checked periodically and, if necessary, re-established so that a cyberattack can be launched when necessary. Limiting the focus of strategic cyber operations to one or a few individual sectors conserves the resources needed for maintaining access

Time-urgent *ad hoc* cyber response options are very difficult if not impossible. By definition, an *ad hoc* operation is one that is not preplanned and detailed intelligence collection on a previously uninvestigated target is time-consuming. Thus, it is unrealistic to expect prompt execution of cyber options against newly identified targets.

Finally, cyber operations whose effects take a long time to manifest are unlikely to have sufficient deterrent effect, a proposition that if true rules out cyber-enabled information operations as a plausible response option.

The remainder of chapter 5 builds upon and elaborates the points made above.

4.2. **Retaliation Means**

Traditional deterrence by punishment has moved from primarily focusing on (the threat) of military action to including other forms of statecraft, most notable diplomatic, information, and economic leavers commonly abbreviated as DIME.[142] After all, conflict is not limited to purely military means, but each instrument of national power has its role both below and above the war threshold.[143] Here, the predominant focus is placed on the information and military domain as it relates to deterrence *through* cyberspace, and its effects on the other domains, rather than the deterrence *of* cyber attacks, which inevitably would have a much larger scope. This does not presuppose that the military domain is the most important lever in deterrence, indeed some scholars might consider other levers to be equally or even more important. After a short introduction of the potential retaliatory means for the other instruments of statecraft, the cyber means of response are parsed into strategic cyber effects and special cyber effects. Then follows the notion of a minimum deterrence capability that allows SMPs to not win the war but to still inflict an unacceptable level of retaliatory punishment on a potential aggressor, no matter their overwhelming technical superiority. The operational considerations of such a capability are evaluated by transposing a Single Integrated Operational Plan (SIOP) that lays out a multifaceted escalation process from an operational point of view to inform military strategic planning. Informed by the psychological and political effects of cyber operations, the targeting options of a traditional SIOP – being counterforce (military) and countervalue (economic or wider societal) – are complemented with a third category: *counterpolitical* targets. To inform the targeting and the intelligence requirements for this category, lessons are drawn from sanction regimes and counterinsurgency respectively.

Diplomatic and economic means are two separate instruments that are increasingly intertwined within the European context when it comes to below-the-threshold retaliation. Diplomatic measures are primarily intended to signal disapproval of the actions of another state and include demarches, bilateral channels, public attribution and naming and shaming, and diplomatic expulsions. States will first try signaling their disapproval privately towards

[142] Brandon Morgan, "Dropping Dimes: Leveraging all Elements of National Power on the Multi-domain battlefield," Modern War Institute, (September 18, 2019)

[143] US Department of the Army, "Joint Doctrine Note 1-8 Strategy," (April 25, 2018), 25.

another state, but when there is a lack of response from the counterparty, they often move towards public attribution. Nonetheless, a case study of the European and US diplomatic response to Russian cyber operations by APT-28, between 2016 and 2018, shows that detailed public attributions have been able to impose costs against aggressors on an exceptional basis.[144] Second, diplomatic expulsions, like the one following the 2016 Presidential interference[145] or the 2020 SolarWinds cyber espionage campaign[146] go one step further in imposing costs and often result in a tit-for-tat move from the other side. This kind of diplomatic retaliation may be damaging to the reputation of the accused state, but is often criticized for falling short when dealing with an actor who simply shrugs off or rejects any accusations on the basis of a lack of evidence, making it largely a symbolic response.

More coercive measures include indictments and economic sanctions. In particular, the US has been actively indicting foreign hackers. Since 2014, the US Department of Justice has released numerous indictments against both individuals found to have acted in close relation with their respective governments or military units, and cybercriminals with no apparent governmental link.[147] They are carried out by law enforcement agencies to target individuals, rather than states, for criminal wrongdoing on the basis of domestic legislation.[148] They also require evidence that meets the requisites of probable cause whereas public state attributions have no evidence thresholds. Such evidence, however, often relies on classified intelligence assets or methods that the respective state would rather not disclose.[149] Economic measures include financial sanctions (asset freezes), trade embargoes (flight and shipping bans, limitations of export and access to markets), foreign assistance reductions and cut-offs, arms embargoes (prohibition of weapon and dual-use exports), and travel restrictions (visa bans). Both within the EU and the US context, sanctions targeting malicious cyber operations are primarily directed at persons or organizations rather than states.[150] In total, the US Department of Treasury has issued more than 140 cyber-related sanctions against Russian

[144] Following the British response to the September 2018 poisoning of Sergei Skripal and subsequent Dutch response to the OPCW operation, a high level of evidence was disclosed, including identities and personal data of the GRU officers they believed to be responsible. This allowed independent investigative collective Bellingcat to expose a major data breach disclosing the identities of approximately 305 GRU officers. Faesen et al, "From Blurred Lines to Red Lines. How Countermeasures and Norms Shape Hybrid Conflict."

[145] BBC News, "US expels Russian diplomats over cyberattack allegations," BBC (December 29, 2016).

[146] Eric Tucker and Aamer Madhani, "US expels Russian diplomats, imposes sanctions for hacking," ABC News (April 16, 2021).

[147] For example, two Ukrainian and Russian nationals were indicted for their role in the Kaseya ransomware attack in 2021, four Chinese PLA officers were indicted for the Equifax hack in 2020, 12 Russian GRU officers were indicted for hacking the DNC and DCCC networks, and North Korean military hackers were indicted for their role in the Sony cyberattack and WannaCry ransomware in 2018, to name just a few. John Sakellaridis, "How the Justice Department Is Stepping up Its Efforts To Indict State-Sponsored Hackers," The Record (February 3, 2021).

[148] In the US case, the most cited legal basis for the indictments concerning malicious cyber operations derive from the Computer Fraud and Abuse Act: Charles Doyle, "Cybercrime: An Overview of the Federal Computer Fraud and Abuse Statute and Related Federal Criminal Laws", Congressional Research Service, (15 October, 2014). Carrie Johnson, "U.S. Charges 7 Russian Intelligence Officers With Hacking 40 Sports And Doping Groups", NPR (2018), US Department of Justice, "U.S. Charges Russian GRU Officers With International Hacking and Related Influence and Disinformation Operations." (October 4, 2018).

[149] When Concord, a Russian troll company, charged by the US Mueller indictment, contested the charges brought against it, the US Justice Department dropped the charges to preserve national security interests and prevent Russia from weaponizing lawful protocols to acquire delicate American law enforcement information. Katie Benner and Sharon, LaFraniere, "Justice Dept. Moves to Drop Charges Against Russian Firms Filed by Mueller", New York Times, (2020).

[150] Jason Bartlett and Megan Ophel, "Sanctions by the Numbers: Spotlight on Cyber Sanctions," CNAS, (May 4, 2021). In the US, the Treasury Department is the agency and does so based on Executive Order 13757 and 13694 that specifically deal with cyber-enabled activities, as well as pre-existing sanction statutes and regulations. The Russian operatives sanctioned by the US were done pursuant to the Countering America's Adversaries Through Sanctions Act (CAATSA). US Department of the Treasury, "Sanctions Related to Significant Malicious Cyber-Enabled Activities" (2019).

nationals[151], including the sanctioning of 32 entities and individuals for their role in the 2020 attempted election interference.[152] Since June 2017, the EU established its cyber-related sanction framework through the EU Cyber Diplomacy Toolbox,[153] which it has only used twice thus far: once against officials of the Russian Federation for their role in the 2015 German Bundestag attack,[154] and once in response to Russian, Chinese and North Korean hacks.[155] Using the Toolbox requires unanimity from all EU member states, which may make its use problematic considering some member states' energy entanglement and dependency on Russia.

From a strategic standpoint, the decision to opt for a specific cyber operation over another depends on the preferred end goal, whether this is a circumscribed denial of a service or a major attack disrupting the power grid of a country.

Information and military responses are predominantly part of the cyber active measures explored in the previous chapter, including cyber support operations and cyberattack operations. From a strategic standpoint, the decision to opt for a specific cyber operation over another depends on the preferred end goal, whether this is a circumscribed denial of a service or a major attack disrupting the power grid of a country. This consideration informs the overarching goal of deterrence through cyberspace, whereby cyber operations can be employed by the deterrer to make the aggressor uncertain of its ability to achieve its goals rapidly. This uncertainty could potentially deter the aggressor from either taking action in the first place, or from escalating it further.

Based on the taxonomy, cyber attack operations constitute the most obvious retaliation means, but other measures can be used as well. An offensive counter cyber operations can be used to target a botnet server that has forced a country to shut down all non-essential Internet traffic, or to insert malware against adversarial military infrastructure, such as the telecommunications system. Cyber support operations can include cyber and electromagnetic activities or counter-information warfare activities, such as actively targeting terrorist recruitment and propaganda efforts, as displayed in US operations (including Glowing Symphony) against ISIS. However, most deterrence through cyberspace operations will focus on the element of 'cyber attack'. These operations can take place at the tactical and strategic level, targeting both *counterforce* (military) and *countervalue* (civilian) targets. The former includes adversarial command and control infrastructure, while the latter includes operations against critical infrastructure such as power grids,[156] petrochemical plants,[157] energy systems,[158] water dams,[159] water systems,[160] and gas pipelines.[161] They include a wide range of activities around or beyond the use of the force threshold that is executed via Internet technologies, or even the Internet itself. These can be exercised with *strategic* or *special* effects, whereby

[151] Jason Bartlett and Megan Ophel, "Sanctions by the Numbers: Spotlight on Cyber Sanctions,"

[152] Carrie Mihalcik, "US sanctions Russia over SolarWinds hack, election interference," cnet.com, (April 15, 2021).

[153] Council of the European Union, "Council Decision (CFSP) 2020/1537," EUR-LEX Document 32020D1537 (22 October 2020).

[154] Council of the European Union: "Council Regulation (EU) 2019/796 of 17 May 2019 Concerning Restrictive Measures Against Cyber-Attacks Threatening the Union or its Member States" EUR-LEX Document 32019R0796, (2019)/

[155] European Council, "EU Imposes the First Ever Sanctions against Cyber-Attacks", (30 July, 2020). Council of the European Union, "Council Implementing Regulation (EU) 2020/1125," EUR-LEX (July 30, 2020).

[156] Zetter, "Inside the Cunning, Unprecedented Hack of Ukraine's Power Grid,"

[157] U.S. Department of the Treasury, "Treasury Sanctions Russian Government Research Institution Connected to the Triton Malware," *U.S. Department of the Treasury,* (October 23, 2020)

[158] Sean Lyngaas, "Taiwan's state-owned energy company suffers ransomware attack," *CyberScoop,* (May 5, 2020) https://www.cyberscoop.com/cpc-corp-ransomware-attack-taiwan-trend-micro/.

[159] Gary Cohen, "Throwback Attack: How the modest Bowman Avenue Dam became the target of Iranian hackers," *Industrial Cybersecurity Pulse,* (August 12, 2021).

[160] Pierluigi Paganini, "Piping botnet: Researchers warns of possible cyberattacks against urban water services," *Security Affairs,* (August 16, 2018).

[161] Christina Wilkie, "Colonial Pipeline paid $5 million ransom one day after cyberattack, CEO tells Senate," *CNBC,* (June 8, 2021).

the former category represents a wholly new instrument for most small and medium powers. Based on the US experience, the rationale is often to subdivide these effects based on the legal authorities that they depend on – Title 10 or Title 50 under US Code for armed forces and national defense (espionage), respectively. In many liberal democracies a similar subdivision has taken place. In the Netherlands, the intelligence law does not allow for cyberattack operations, with the exception of a limited number of countermeasures.[162] Such operations are executed by the armed forces according to the mandate in Article 97 of the Constitution, which includes maintaining or promoting the international legal order, and according to the process described in Article 100.[163]

However, the principal difference between strategic and special effects is not only one in legal authorities, but also in their bearings on international law. Armed forces are invariably bound by International Humanitarian Law (IHL) and must also worry about the precedents of certain actions. Intelligence special operations will also abide by international law, but perhaps less stringently, and can exploit legal gray zones as they will usually not be made public. In the view of the authors, the primary difference between strategic and special cyber effects is a combination of factors – primarily the target, the effects, and the overall conflict context. A one-off target that is hit without rising above the threshold of armed attack in a time of relative peace is more likely to be considered as an intelligence special operation. Indeed, as Klimburg has argued, there is a strong case to say that most one-off cyber activities that take place in a legal gray area – for instance, the USCYBERCOM disruption of the Russian IRA troll factory – should not be exercised overtly under military authorities (or even widely advertised) in order to prevent negative precedents being set. The concept of special operations allows more offensive activity to take place without officially condoning it – even if often the type of activity that can take place could also easily be described as 'strategic cyber' in another conflict context. Therefore, wherever possible, cyberattack options should be considered under these authorities before being branded as regular strategic cyber operations. There are three exceptions to the rule of 'special' before 'strategic'. Again, the determinants should be scale, impact, and context. Any large cyber operation that is clearly multi-pronged, involves different entities, has cumulative effects that are above the armed attack threshold, and takes place in a state of belligerency should be considered a regular operation of the armed forces and therefore, a strategic cyber attack fully bound by the Laws of Armed Conflict.

The primary difference between strategic and special cyber effects is a combination of factors – primarily the target, the effects, and the overall conflict context.

4.3. **SIOP Minimum Deterrence?**

As a US DoD study indicated in 2013, offensive cyber can both be very expensive and very cheap at the same time (see Annex III for more information on the various tiers of offensive capability).[164] At the highest levels, cyber campaigns can consume many tens of millions of euros or more to prepare, involve thousands of manhours of cutting-edge bespoke coding,

[162] "Wet op de inlichtingen- en veiligheidsdiensten," *Algemene Inlichtingen- en Veiligheidsdienst. Ministerie van Binnenlandse Zaken en Koninkrijkrelaties.*

[163] Once the Cabinet decides on a military operation, it usually informs parliament directly and usually in advance. The procedure varies when cyber operations are categorized as special operations. In this case, the decision is taken by the Ministerial Core Group on Special Operations, which also decides when to inform parliament, which can occur after the mission in case of "substantial political and military risks and the need for strict secrecy."P.A.L Ducheine, K.L Arnold, and B.M.J Pijpers, "Decision-Making and Parliamentary Control for International Cyber Operations by the Netherlands Armed Forces," *Amsterdam Center for International Law,* (2020), 16-17.

[164] Depart of Defense, Defense Science Board, "Task Force Report: Resilient Military Systems and the Advanced Cyber Threat," (January 2013), Office of the under Secretary of Defense for Acquisition, Technology and Logistics, Washington, D.C. See also Klimburg*, The Darkening Web*, 2018.

in-depth reconnaissance, testing (even using replicas of physical equipment or networks), and not least the use of a number of zero-day (previously unknown) exploits. Furthermore, they can require special means of inserting the code into the targeted networks, for instance by human agents, which has its own considerable costs attached. Finally, a large part of these costs is reoccurring. As networks change and programs get patched, it is often necessary to repeat most components of the attack chain. Depending on the target, this could easily occur many times a year. In public research two examples of such high-end cyber attacks have been referenced, the first being the well-known Stuxnet (part of the Olympic Games Operation) attack on Iran, the second being the purported in-depth US Nitro Zeus campaign targeting the Iranian defense infrastructure.[165]

The objective of minimum deterrence is not to win a war but to inflict unacceptable costs to another actor and prevent them from winning without major costs.

However, the vast majority of cyberattacks can be much less advanced and still be very disruptive or even destructive – if not at the same targets. The example here is Shamoon, the purported response by the Iranians to Stuxnet. Using repurposed US malware (the 'Wiper' module also found in Flame), the cyberattack hit Saudi ARAMCO, one of the largest oil companies in the world, and caused significant damage. With a little more work, the damage could have been catastrophic to the company, and likely have had a significant impact on its operational performance. This additional effort would not have been hard to deliver – the original Shamoon attack was immeasurably cheaper (maybe even to the factor of 100) than Stuxnet. Not only did it have a physical effect, but to this day Iran can claim to possess a sort of cyber deterrence capability for just a fraction of the cost of the US.

The cyber punishment potential of small-to-medium cyber powers may be orders of magnitude less than that of the US or a near-peer cyber power, but some of these countries still possess a 'minimum deterrence capability' that, much like the small nuclear arsenals of France, the UK and China, could inflict an unacceptable level of retaliatory punishment on a potential aggressor, no matter their overwhelming technical superiority. These nations may possess something which was previously unavailable to them: not just a strategic weapons capability – a virtual strike force no less potent than a wing of bombers or ballistic missiles – but also a defensive advantage towards larger foes. The sheer number of nations that may be able to compete with and reciprocally threaten a major power could be historically unprecedented. From the perspective of these SMPs, understanding how to integrate this asymmetric advantage of cyber deterrence into a broader operational framework and project it into a Whole of Nation/Union/Alliance deterrence posture is crucial. Turning to the integration of such capabilities within the overall deterrence posture, later on, the remainder of this chapter explores ways to integrate a 'minimum deterrence capability' within the operational framework as part of a 'Single Integrated Operational Plan' (SIOP) for cyber.

The objective of minimum deterrence is not to win a war but to inflict unacceptable costs to another actor and prevent them from winning without major costs. This does not mean that such deterrence merely rests on the certainty of inflicting unacceptable damage on an aggressor, but on the potential aggressor's uncertainty of avoiding unacceptable damage.[166] One of the earliest definitions of the term calls it "an attempt to prevent enemy attack through reliance on a small nuclear retaliatory force capable of destroying a limited number of key

[165] David E. Sanger and Mark Mazzetti, "U.S. Had Cyberattack Plan if Iran Nuclear Dispute Led to Conflict," New York Times, (February 16, 2016), https://www.nytimes.com/2016/02/17/world/middleeast/us-had-cyberattack-planned-if-iran-nuclear-negotiations-failed.html

[166] Rajesh M. Basrur, "Minimum deterrence: Fundamentals and policy implications," *Indian Foreign Affairs Journal, vol. 1* no. 3, (2006), 63.

targets."[167] During the Cold War, the arms race between the US and the Soviet Union drove China towards the pursuit of a minimum deterrence strategy, which aimed to develop a nuclear arsenal that was substantial enough to demolish adversarial strategic infrastructure under a 'no first use' policy.[168] Both India and Pakistan have adopted similar policies to justify their nuclear stockpiling. Inflicting unacceptable damage through cyberspace depends on the political objectives of the potential aggressor. Some countries may find temporary shutdowns of critical infrastructure unacceptable, while others may consider inciting their own regime security as sacrosanct. As described by Klimburg, in contrast to the Cold War, where both West and East shared the common nightmare of the nuclear mushroom cloud, they currently hold different nightmares. While the West fears cyber operations that paralyze its critical infrastructure, the so-called East has an additional, arguably more important, nightmare of cyber operations, namely that of information warfare undermining its regime's security.

The rise of a minimum strategic cyber deterrent capability offers small nations the new ability to strike an adversary in their strategic depth – hitting military supply chains away from the forward edge of the battle area, and including critical infrastructure, top-level command and control, and even sources of political and economic power. For nations that did not previously have such a strategic military capability, this development represents the single greatest paradigm shift at least in recent history. Suddenly, it is possible for a smaller state to reciprocally threaten a much larger rival directly. This is a major conceptual challenge for both orthodoxly trained military thought as well as political leadership. Further, below the threshold of declared war, the increased prevalence of special cyber effects has increasingly become a matter of national concern, with espionage, sabotage, and active information warfare operations increasingly becoming a reality. Although smaller democratic nations are more often victims than perpetrators of these actions, they too can consider a number of new offensive options as another form of deterrence and retribution. The range of special effects technically possible with cyber means is nearly unlimited. From more classic special operations to covert intelligence and fully novel forms of information warfare, cyber opens up a wide range of options but also threats that need to be countered.

The question of whether there is a need for a Single Integrated Operational Plan (SIOP) for cyber, and what it would look like, has been raised by Austin Long, who focuses on the operational considerations for strategic offensive cyber military planning.[169] Long argues that discussions about offensive cyber operations have thus far mirrored the early nuclear age, often neglecting command, control, communications and intelligence (C3I) and operational issues in favor of more theoretical debates. Drawing from the US lexicon on nuclear war planning, he aims to pull the discussion on strategic offensive cyber operations (OCO) towards the operational and tactical considerations. Overall, he finds that OCO planners should receive guidance from the National Command Authority (NCA – i.e. the White House) on how to structure a cyber offensive plan. In particular, he finds that NCA should provide further clarity on the objectives for more accurate OCO targeting, on the attack structure and operational priorities, and should increase understanding of OCO damage expectancy.

[167] John Baylis and Ken Booth, "*Contemporary strategy: theories and policies*," New York: Holmes & Meier, (1987), 312

[168] Their recent nuclear build-up in the shape of the expansion of atomic missile silo in North-Central China proves a turn-around in their policy that is consistent with the overall military budget increases over the past decades. See: Hans M. Kristensen and Matt Korda, "China's nuclear missile silo expansion: From minimum deterrence, to medium deterrence" The Bulletin, (September 1, 2021).

[169] Austin Long, "A Cyber SIOP? Operational considerations for strategic offensive cyber planning," *Journal of Cybersecurity, vol. 3*, no. 1, (March 2017), 19,20.

Transposing the nuclear SIOP to cyber allows Long to parse the main categories of targets for offensive cyber operations and the nature of effects they could generate on those targets as well as their intelligence and wider C3I requirements. In terms of targeting, a distinction can be made between two categories, namely countervalue and counterforce targets. Counterforce refers to targets that have a significant military utility, such as the Flame campaign aimed at Iran's Revolutionary Guard Corp operating outside of Iran, or part of Russia's operations in Ukraine and Georgia, or the Israeli Operation Orchard that allowed the Suppression of Enemy Air Defenses. Countervalue targets are non-military targets that have a wider economic or societal value but remain strategically relevant, such as industrial, energy or financial targets. Examples include a wide range of cyber operations, including BlackEnergy or Stuxnet. While countervalue targets are predominantly chosen with a deterrence by punishment strategy in mind as they drastically increase the perceived costs of aggression, counterforce targets were considered for both deterrence by punishment and denial. After all, holding adversarial military capabilities at risk not only imposes costs but can also appear to have degrading effects on said capabilities, possibly to such an extent that they would not achieve their military objectives.

Striking such a balance – exercising sufficient resolve without triggering escalation – will be a major challenge, in no small part because of the risks of inadvertent escalation.

It should be noted that counterforce and countervalue targets partly overlap depending on the context and the overall objective. A countervalue target, like an electricity or gas supplier, can become a counterforce target as soon as military assets rely on it. This overlap is much deeper in cyberspace – an environment known for its dual-use technology and infrastructure. While some systems may be exclusively and uniquely reserved to the military, they often rely on commercial products or operate on infrastructure that is shared with civilian users. At the same time, compared to conventional countervalue means, where the impact is geographically limited to the selected target, cyber means can effectively paralyze a specific asset of a whole nation.[170] While this brings important strategic value to the deterrer, it can, as previously explained, increase the risk of second-order effects and the risk of escalation.

Based on these two categories, Long considers a range of targeting sets and SIOP components. In terms of the overall objectives, deterrence through the threat of retaliation is the first obvious parallel. Second, when deterrence fails, the plan seeks to establish escalation control by limiting the scope of response and specific targeting. Third, when escalation control fails, it sets out to engage in a war to achieve maximum power. Transposing these three steps to the cyber context, we first have to look at what we are trying to deter through cyber operations. Given the focus on deterrence through cyberspace, they can be part of a broader set of deterrent capabilities, but can also be used to deter adversary cyber operations. If deterrence then fails, the planners need subsequent objectives. Following the SIOP analogy, it could very well include escalation control or the neutralization of adversary strategic capabilities in order to limit damage from enemy attacks depending on the opponent. Cyber operations can contribute to either option but would have to be planned differently according to their escalatory effects. If the objective is escalation control, the focus should be on targets that produce sufficient punishment or denial effects but that are unlikely to lead to further escalation by the adversary. Striking such a balance – exercising sufficient resolve without triggering escalation – will be a major challenge, in no small part because of the risks of inadvertent escalation.

Finally, it should be noted that deterrence continues also when peacetime deterrence fails. The objective may then be primarily driven by operational considerations, rather than political or psychological messaging. It is, however, a mistake to let operational considerations take precedence over a single overriding strategic narrative. The following section will show

[170] Smeets, "The Strategic Promise of Offensive Cyber Operations" Strategic Studies Quarterly, (Fall, 2018), 99.

how cyber operations have considerable political and psychological effects that need to be considered equally next to the operational considerations.

4.4. Counterforce, Countervalue, and... Counterpolitical?

The SIOP for cyber lays out a multifaced escalation process from an operational point of view to inform military strategic planning. However, the national priorities of states in the cyber context dictate that minimum deterrence targeting should extend beyond the purely military and extend to the political. In other words, effecting maximum political effect on the opponent is crucial when it comes to minimum deterrence. This introduces a third category of *counterpolitical* targets that are an addition to the strategic retaliatory capacity for minimum deterrence. In contrast to countervalue targets, these targets would have no sweeping effects on the national economy. Instead, they are narrowed down to specific targets of high political value, such as individual oligarchs or companies that hold high intrinsic and psychological value within a country, even if the value is not widely shared or completely hidden from public view. Such an approach accepts the information warfare narrative that is being pushed by some states and which undergirds their overall political objectives and strategy, but refuses its means – namely, the open media ecosystem. Instead, it can be compared to the approach adopted in certain economic sanction regimes, where targets are selected based on political effects (e.g. sanctioning an industry located in a nation's region that holds an important political role, for example in upcoming elections).

Lessons can be drawn from the counterinsurgency discourse, primarily the (ASCOPE) PMESII matrix, as a basis for targeting in any concept of operations.

In a scenario of increasing escalation, clear strategic objectives, a structured attack plan, strong consideration of potential damage, and advance planning are fundamental to envision the use of offensive cyber capabilities in a way that fits a strategy of deterrence. Thoroughly tackling all these points can facilitate their employment. Success largely depends on the length of time and level of preparation required, often involving advance planning and pre-deployment. This will allow the deterrer to enhance the level of precision of the target and limit unintended effects. Thorough preparation further allows the actor to keep control over the escalation ladder and have the ability to de-escalate the conflict, once it has gained leverage over its adversary. Traditionally, such intelligence needs and preparation requirements occur mostly on the operational level to create an operational picture of the target, without much attention to the political, economic or other considerations that can contribute to determining the political value of a target. To fill this gap, lessons can be drawn from the counterinsurgency discourse, primarily the (ASCOPE) PMESII matrix, as a basis for targeting in any concept of operations.[171] It contributes to a holistic understanding of the operational environment of friendly, neutral, and threat political military, economic, social information, and infrastructure (PMESII) systems – "a set of interrelated operational variables that provides counterinsurgents with a method to analyze the operational environment through specific filters".[172]

[171] US Joint Publication 3-24: Counterinsurgency (April 2021). The PMESII framework is also recognized by the Dutch Ministry of Defense in "Joint Doctrine Publicatie 5 Commandovoering" and "Landoperaties: Doctrine Publicatie 3.2".

[172] Counterinsurgency Training Center Afghanistan, "A Counterinsurgent's Guidebook," Camp Julien, Kabul, Afghanistan, (Version 2: November 2011).

Figure 2. ASCOPE/PMESII Framework. Derived From US Marines, "Planning Templates October 2017".

	P	M	E	S	I	I
	Political	Military	Economic	Social	Information	Infrastructure
A Areas	Areas – Political (District Boundary, Party affiliation areas)	Areas – Military (Coalition / LN bases, historic ambush/IED sites)	Areas – Economic (bazaars, shops, markets)	Areas – Social (parks and other meeting areas)	Areas – Information (Radio / TV / newspapers, where people gather for word-of-mouth)	Areas – Infrastructure (Irrigation networks, water tables, medical coverage)
S Structures	Structures – Political (town halls, government offices)	Structures – Military / Police (police HQ, Military HHQ locations	Structures – Economic (banks, markets, storage, facilities)	Structures – Social (Churches, restaurants, bars, etc.)	Structures – Information (Cell / Radio / TV towers, print shops)	Structures – Infrastructure (roads, bridges, power lines, walls, dams)
C Capabilities	Capabilities – Political (Dispute resolution, Insurgent capabilities)	Capabilities – Military (security posture, strengths and weaknesses)	Capabilities – Economic (access to banks, ability to withstand natural disasters)	Capabilities – Social (Strength of local & national ties)	Capabilities – Information (Literacy rate, availability of media / phone service)	Capabilities – Infrastructure (Ability to build / maintain roads, walls, dams)
O Organizations	Organizations – Political (Political parties and other power brokers, UN,)	Organizations – Military (What units of military, police, insurgent are present)	Organizations – Economic (Banks, large land holders, big businesses)	Organizations – Social (tribes, clans, families, youth groups, NGOs / IGOs)	Organizations – Information (NEWS groups, influential people who pass word)	Organizations – Infrastructure (Government ministries, construction companies)
P People	People – Political (Governors, councils, elders)	People – Military (Leaders from coalition, LN and insurgent forces)	People – Economic (Bankers, landholders, merchants)	People – Social (Religious leaders, influential families)	People – Information (Media owners, mullahs, heads of powerful families)	People – Infrastructure (Builders, contractors, development councils)
E Events	Events – Political (elections, council meetings)	Events – Military (lethal/nonlethal events, loss of leadership, operations, anniversaries)	Events – Economic (drought, harvest, business open/ close)	Events – Social (holidays, weddings, religious days)	Events – Information (IO campaigns, project openings, CIVCAS events)	Events – Infrastructure (road / bridge construction, well digging, scheduled maintenance)

While all military campaigns since WWII will lay claim to some understanding of the interlocking nature of the enemy leadership, physical infrastructure and political networks, the proposed emphasis goes well beyond standard attempts by known military thinkers (e.g. John Warden's Five Rings).[173] It also draws heavily on the Effects-Based Approach to Operations

[173] Conceived by Colonal John Warden, the Five Rings model was developed to provide "valuable guidance in breaking down an enemy into a system, thereby dissecting the critical nodes with the goal of identifying centers of gravity (COGs)". The model consists of (1) leadership, (2) systems essentials, (3) country infrastructure, (4) population, and (5) fielded forces. For more information, see Russell J. Smith, "Developing an Air Campaign Strategy," *Air University,* https://www.airuniversity.af.edu/Portals/10/ASPJ/journals/Chronicles/smith.pdf

(EBAO) concept, the popularity of which has waxed and waned in recent decades.[174] There is no doubt that in the operational context even Western cyberattacks have accepted the need to concentrate on political effects as much as any operational effects. Smeets already highlighted the significant advantage brought by the employment of cyber means in a nation's deterrence strategy is the application of counterforce and countervalue to the cognitive domain. Offensive cyber operations have strong psychological consequences of "humiliation and confidence degradation" linked to the exposure of vulnerabilities.[175] This objective, coupled with the general desire of the targeted state to save face, is pursued both for cyber countervalue and counterforce targets. For example, one of the goals of the Stuxnet cyber-attack was to embarrass the Iranian government. Similarly, Operation Orchard is generally considered a humiliation for Syrian President Assad.[176] As previously argued, some nations even rank the psychological or political effect of certain cyber operations significantly higher than the possible operational benefits. The prime example here is potentially NotPetya ransomware campaigns, which caused, seemingly incidentally to the intended targets in Ukraine, over 10 billion dollars' worth of damage to multinational companies in other countries. The intended effects here may well have been a political warning shot: an attempt to feel out the will of the US and UK; an attempt to influence the wider political narrative of cybersecurity; or indeed all of the above. In any case, the emphasis on a *counterpolitical* objective is a given. And arguably this operation could be classified as a *special* operation – not strategic, yet another indication that what happens in peacetime circumstances can be expected in wartime as well. Likewise, a SMPs'cyber deterrent should be able to respond and ideally deter such activity from occurring. This can include conducting a similar counterpolitical strike – even by using other targets and means.

Offensive cyber operations have strong psychological consequences of "humiliation and confidence degradation" linked to the exposure of vulnerabilities.

4.5. **Main Takeaways**

Deterrence by punishment has shifted from solely military to encompass diplomatic, informational, military and economic punitive tools. Designed to signal disapproval, diplomatic actions, such as public attribution and diplomatic expulsions, have been increasingly used by states. Although there have been instances where collective diplomatic response led to the imposition of costs on the aggressors, they have largely remained symbolic. Economic measures, such as sanctions, are regarded as more coercive. Their effectiveness has been questioned although in certain instances, such as the US response against Chinese IP theft, when they were taken in conjunction with diplomatic measures, they affected the cost-benefit calculus of the opponent in an effective, albeit short-lived, way.

Deterrence though cyberspace, offers small-to-medium states an unprecedented form of retaliation – or a minimum deterrent capability that can hit the adversary in peacetime and outside of the battlefield. Such retaliation can be separated into two forms: the strategic cyber effects versus the special cyber effects. This subdivision is often based on the legal authorities they depend on (strategic effects by the armed forces and special effects by the intelligence community) and has bearings on international law. The latter category is less constrained by the bounds of international law and given its covert nature, there are fewer concerns over the precedents of certain actions. Beyond the legal mandates, the primary

[174] See Air Force Doctrine Publication (AFDP) 3-0 Operations and Planning, "The Effects-Based Approach to Operations (EBAO)," *Curtis E. Lemay Center,* (November 2016). https://www.doctrine.af.mil/Portals/61/documents/AFDP_3-0/3-0-D06-OPS-EBAO.pdf

[175] Smeets, "The Strategic Promise of Offensive Cyber Operations", 101

[176] Ibid., 102

difference between these two effects is, in our view, a combination of the target, the effects, and the overall conflict context. A one-off operation in relative peace below the armed attack threshold is more likely to be considered an intelligence special operation. A multi-pronged offensive campaign with cumulative effects that moves beyond the armed attack threshold or takes place in a state of belligerency should be considered a regular operation of the armed forces and therefore, a strategic cyber attack. While covert responses only have a direct deterrent effect on the target–in contrast to overt responses that can also deter other potential adversaries–the concept of 'special operations' allows much offensive activity to take place without officially condoning it, thereby avoiding risky precedents.

The cyber punishment potential of SMPs may be orders of magnitude less than that of the US or its near-peer cyber power, but some of these countries still possess a minimum deterrence capability that, much like the small nuclear arsenals of France, the UK and China, could inflict an unacceptable level of retaliatory punishment on a potential aggressor. The value of minimum deterrence is not to win the war, but to raise the perception that one can inflict unacceptable costs to another actor. Some of these nations may lack the resources of the higher tier countries to achieve timely retaliatory effects across a wide range of countervalue and counterforce sectors. Depending on the strength and robustness of the target's defenses, they will have to concentrate their resources on fewer individual countervalue sectors to achieve significant effect rather than distributing them over a large number of sectors. Countervalue targets are more likely to have poorer cyber defenses than counterforce targets and crippling one critical sector will have great effect on an adversary than modest damage across many sectors. This also conserves intelligence resources, which are significantly more limited for small and medium powers compared to larger powers. Nonetheless, an intelligence lift is required to extend beyond the standard operational approach and include psychological, political, economic and other considerations. To this end, the counterinsurgency (ASCOPE-)PMESII framework can function as a useful tool to determine the intelligence needs required for better understanding of the wide operational environment that can contribute to determining the political value of a target. Success will still depend on the length of time and level of preparation required, which often involves advance planning and pre-deployment. It is therefore unrealistic to expect a prompt execution of cyber operations against newly acquired targets. Finally, a third category of targets is introduced: counterpolitical targets. In contrast to other targets, they do not have sweeping effects on a nation's military or economy. But, in a similar vein to some sanction regimes, these targets hold high intrinsic and psychologic value within a country, even if the value is completely hidden from public view and not widely shared. They offer an additional avenue for covert punishment (often below the war threshold) that effectively strikes the opponent without them noticing.

5. Retaliation paths and organizational considerations

5.1. Introduction by Erica Lonergan

Cyber deterrence presents vexing challenges for policymakers due to several factors, such as the challenges associated with attribution, low barriers to entry, the multitude of actors conducting malicious behavior in cyberspace, limitations of demonstrating retaliatory capabilities via cyber means, and so on.[177] However, across all of these issues, a critical factor that complicates cyber deterrence is developing approaches that meaningfully integrate the diverse stakeholders that play a role in ensuring effective defenses, resilience, and responses, given the interdependent and transnational nature of the domain.

It has become a truism to describe cyberspace as a multistakeholder environment. Yet, this chapter aims to move beyond this fundamental assessment to identify how different groups of stakeholders grapple with working together to develop and implement effective deterrence strategies. This spans the various agencies and entities within a particular government, the private sector, and international partnerships. Specifically, from a Whole of Government perspective it involves information-sharing and coordination across the various arms of government that are responsible for national cybersecurity, which includes entities beyond the traditional national security and intelligence organizations. From a Whole of Nation or Whole-of-Society perspective it entails implementing collaborative models between the government and the private sector—especially the owners and operators of critical infrastructure—to defend a nation in cyberspace, as well as incorporating everyday citizens who play a role in the cybersecurity ecosystem. Finally, from a Whole of System perspective it includes determining how allies and partners can cooperate to achieve shared strategic objectives, despite considerable heterogeneity among allies in terms of capabilities, willingness to conduct offensive cyber operations, conceptualization of gray and red space, definitions of red lines, and other key matters.

Governments do recognize the challenges of multistakeholder models for cyber deterrence. For example, the Cyberspace Solarium Commission, chartered by the US Congress in the 2019 National Defense Authorization Act, advances a new strategic approach of "layered cyber deterrence."[178] A central premise of this strategy is that the US government cannot deter cyber attacks in the absence of significant collaboration with the private sector, with a focus on systemically important critical infrastructure. A number of core recommendations

[177] Erica D. Borghard and Shawn W. Lonergan, "Deterrence by Denial in Cyberspace," *Journal of Strategic Studies,* (2021).

[178] US Cyberspace Solarium Commission, "Cyberspace Solarium Commission Final Report," (March 2020),

in the Commission's March 2020 final report are oriented around improving Whole of Government and Whole of Nation structures, processes, and capabilities. A landmark recommendation proposed by the Commission and instantiated in law in the 2021 National Defense Authorization Act was the creation of a Senate-confirmed National Cyber Director in the Executive Office of the President, with a corresponding Office of the National Cyber Director, in order to cohere and coordinate the different elements of the Federal government involved in cybersecurity, and serve as a focal point for interacting with the private sector. Additionally, several measures that receive significant emphasis in the report are aimed at enhancing information-sharing, situational awareness, and analysis among government and critical infrastructure stakeholders; conducting joint planning around likely cyber contingencies and routinely exercising plans; and clarifying roles and responsibilities in the event a cyber incident occurs. Similarly, in the UK's new National Cyber Security Strategy, released in December 2021, the document highlights two of the key shifts in the UK's approach. The first is a focus on more comprehensively bringing together the different elements of government and the private sector for a "truly joined up, national strategic approach;" and the second is investing in a Whole-of-Society effort.[179]

Despite the importance of collaboration, as this chapter highlights, there are enduring challenges that often hamper effectively cohering the different stakeholders involved in cyber strategy. At a Whole –of-Government and Nation level, coordinating roles and responsibilities across the government often demands adjudicating competing bureaucratic interests and prerogatives. For instance, law enforcement agencies tend to privilege investigation and prosecution, while homeland security agencies tend to focus on crisis management and incident response, and these prerogatives are often in tension with one another. Moreover, there is significant variation across states in terms of how government conceptualize the appropriate nature of its relationship with the private sector, such as how much coercive power the government is willing to bring to bear to compel action on the part of the private sector (through regulatory or legislative vehicles). Even defining what counts as critical infrastructure—and therefore what parts of the private sector might require a differently-structured relationship with the government on cybersecurity issues—is a contested issue. For example, in the US, there are sixteen critical infrastructure sectors that are defined by an executive order, as well as subsets of those sectors that are so critical to economic and national security that they warrant distinct status.[180] Yet, what these definitions mean in practice and the implications for how critical infrastructure collaborates with the government, remains uncertain and opaque.

At a Whole –of-Government and Nation level, coordinating roles and responsibilities across the government often demands adjudicating competing bureaucratic interests and prerogatives.

At an international level, the challenges are compounded even when the interdependence of cyberspace and a commonality of threats and interests makes the imperative to cooperate highly salient. The transatlantic alliance, broadly construed, has oftentimes struggled to arrive at a consensus around what constitutes acceptable behavior and appropriate responses in cyberspace. More directly pertinent from a deterrence perspective, there are operational limitations to meaningful cooperation between allies about intelligence-sharing around cyber threats and coordinating responses to malicious cyber behavior. The nexus between cyber and intelligence operations, and the deep ties between the cyber and signals intelligence worlds, creates significant impediments to sharing intelligence even among close allies that would enable timely and effective attribution, defense, and other responses. For offensive cyber operations in particular, the challenge is most delicate. Allies vary in terms of the level of

[179] UK Government, "National Cyber Security Strategy 2022", (15 December 2021).

[180] Obama White House Archive, "Executive Order – Improving Critical Infrastructure Cybersecurity," The White House, Office of the Press Secretary, (February 12, 2013)

maturity of offensive cyber programs (if at all); perspectives on the application of sovereignty to cyberspace and, by extension, how other states should approach maneuvering in others' networks; and potential sources of mistrust about allies' cyber activities. The latter is evident in debates about how new US cyber concepts, such as persistent engagement and defend forward, which were debuted in Cyber Command's 2018 Command Vision and the 2018 Department of Defense Cyber Strategy, might affect US allies.[181] Issues about the extent to which US cyber forces may be maneuvering in allied owned and operated networks, doctrinally defined as "gray space" by the US military, and appropriate mechanisms and timelines for notification about these operations, remain unresolved.[182]

Despite these challenges, there has also been significant progress. Allies have demonstrated greater willingness to conduct joint, public attribution of behavior that violates norms, such as the recent decision by the US, UK, EU, and NATO allies to call out China for the Microsoft Exchange hack.[183] At the NATO summit in Brussels in June, the alliance reaffirmed the applicability of the mutual defense commitment to cyber attacks.[184] And, despite tensions stemming from America's cyber posture, the Defense Department has conducted several dozen "hunt forward" cyber operations that entail US cyber protection teams working with allies and partners to identify and thwart adversary activity on allied networks.[185] As the threat environment continues to pose cyber risks, domestic and international pressure may drive states to take necessary measures to overcome the critical gaps identified in this chapter.

Cyber deterrence is too multifaceted an issue to be dealt with by just the military. A wider range of government mandates play a role, ranging from military and intelligence, to diplomacy and law enforcement.

5.2. **Whole of What?**

Cyber deterrence is too multifaceted an issue to be dealt with by just the military. A wider range of government mandates play a role, ranging from military and intelligence, to diplomacy and law enforcement. The esoteric nature of the many individual mandates involved in cyber deterrence naturally leads to 'stovepiping' in narrowly defined government organizations. The reality of these different mandates is that they are each dealt with by different organizational groups within government, but also within the non-state sector both nationally and internationally. Focusing on the need for different actors to work together on a wide range of interlinked issues, a Whole of Government (WoG) approach is required to improve coordination and unity in action among government agencies. At the same time, they need to be able to operate with international partners (Whole of System) and their national civil society and industry stakeholders (Whole of Nation). Each stakeholder is not necessarily constrained within each category but can operate with multiple 'hats'. This chapter explains how different– government and non-state actors – can interact in cyber deterrence, how they can mutually reinforce each other, and finally what their organizational requirements are for a small-to-medium sized nation.

[181] US Cyber Command, "Achieve and Maintain Cyberspace Superiority"; "Summary: Department of Defense Cyber Strategy," U.S. Department of Defense, "Achieve and Maintain Cyberspace Superiority"; Smeets, "U.S. cyber strategy of persistent engagement & defend forward: implications for the alliance and intelligence collection"

[182] US Department of Defense, "Joint Publication 3-12: Cyberspace Operations".

[183] The White House, "The United States, Joined by Allies and Partners, Attributes Malicious Cyber Activity and Irresponsible State Behavior to the People's Republic of China," The White House, (July 19, 2021)

[184] "Brussels Summit Communique," North Atlantic Treaty Organization, June 14, 2021.

[185] Brad D. Williams, "CYBERCOM has conducted 'hunt forward' ops in 14 countries, deputy says," *Breaking Defense*, November 10, 2021.

5.3. **Whole of Government**

The national Whole-of-Government (WoG) approach, originally a cost-saving method to encourage departments to pool resources and deliver 'more of the same', depends on successful *coordination* between government agencies at the central, state and local level. The most prevalent example is the 3D Approach, consisting of diplomacy, development and defense that is used by many liberal democracies.

Within cyber deterrence, the notion of WoG is the predominant and most important approach, but also the most difficult to achieve. Even for those where the WoG is the stated norm, for example in the US, deterrence efforts, and cyber operations more broadly, are more often than not executed in a closed silo with limited view on outside equities.[186] Of particular importance is the ability for governments to attack and defend in cyberspace and share operational resources, also during a major cyber incident. A WoG approach is effectively achieved when "a unity of purpose across the different level and types of government" is developed and "the improved coordination of national efforts" is accomplished.[187] This is only possible, however, if the legal requirements to mandate such control and coordination among the various governmental bodies are put in place.

Beyond the operational and organizational considerations that inform the crisis management and resilience components of deterrence, there is a strong need for governments to synchronize signaling and public communication at the political, strategic, and tactical level. For example, in response to Chinese economic espionage, the US opted for coercive measures through indictments at the tactical level and the threat of sanctions at the strategic level, while exerting high-level political engagement between Obama and Xi that led to a bilateral agreement. While it operated across different domains and at various levels, Washington signaled consistently and uniformly to Beijing that cyber-enabled IP theft was unacceptable, and that the US was willing to escalate the issue while at the same time offering incentives. This approach not only provided multiple avenues for reinforcement, but it also contained the risk of inadvertent second-order effects, even when overt moves were employed. In contrast, the public communication component of the US persistent engagement doctrine employs a volatile mix of covert military effects and the overt disclosure of them that can lead to mixed signaling and a broad range of unintended and undesirable second-order effects. Several studies have demonstrated that cyber operations in and of themselves fall short when it comes to signaling.[188] Signaling efforts should therefore not just have to rely on the military but should be employed as part of a wider strategic communication effort that coordinates diplomatic and law enforcement signals, ranging from bilateral channels and overt public diplomacy, to indictments and sanctions or other instruments of power of the state and preferably its allies. Deterrence efforts also need to be clearly linked to behavioral benchmarks so the government can effectively communicate why and how it seeks to shape adversary perceptions of the strategic environment, as well as adversary behavior.

[186] Take for example, Michael Haden's reflection on his time as head of the CIA, where he warned about the strategic implications of the US taking down an al-Qaeda website. Allies or partners using the same server may be hit. When he tried to get the military to stop, he was effectively ignored and the tactical mission ended up trumping strategic policy making. As a result, he nearly took the CIA out of the cyber-operations arena. See: Klimburg, *The Darkening Web*, 200 and Michael V. Hayden, "The making of America's cyberweapons," *The Christian Science Monitor,* (February 24, 2016).

[187] Klimburg, "National Cyber Security Framework Manual," 101.

[188] Erica D. Borghard and Shawn W. Lonergan, "The Logic of Coercion in Cyberspace," *Security Studies, vol. 26*, no. 3 (May 2017): 452-481; Brandon Valeriano, Benjamin Jensen and Ryan C. Maness, "Cyber Strategy: The Evolving Character of Power and Coercion" (New York: Oxford University Press, 2018).

One step towards a better Whole of Government approach would be for SMPs to establish a National Security Council. In the Netherlands, traditionally characterized by the polder-model of governance with minimal central direction from the Cabinet's office, such a Council is lacking. But the proposal is not new. The first attempt can be traced back to September 2001, in the aftermath of 9/11, and a motion proposing this idea was actually adopted in 2004, but never implemented. It resurfaced in the following years. In 2016, the Hague Centre for Strategic Studies and Clingendael recommended the establishment of a National Security Council, which was further elaborated on in the 2020 report by the Netherlands Scientific Council for Government Policy.[189] Chaired by the prime minister, it could comprise relevant ministers and the highest-ranking relevant civil servants as well as national experts and representatives of crucial interest groups. Drawing on the outcome of informal meetings that already take place, but which lack a clear constitutional mandate, such a council would lead to a better alignment of domestic and foreign security policy and have a strong link with the research community that complements the analyses of the intelligence and security services through a Security Planning and Research Agency. This agency should be seen as a consortium of existing research institutions that not only informs the Council but also connects it to the wider academic community. It can contribute to an expanded intelligence assessment of the Council and wider buy-in from domestic NGOs for government decisions.

One step towards a better Whole of Government approach would be for SMPs to establish a National Security Council.

Another way for a better Whole-of-Government approach is through a Whole-of-Nation and Whole-of-System strategy, whereby the latter two can help improve the 'culture' and provide additional channels for doing the first one well.

5.4. **Whole of Nation**

Whole-of-Nation (WoN) or Whole-of-Society (WoS) cyber deterrence includes analyzing concepts for a wider defense informatics base to support military cyber operations, the importance of a coherent and transparent coordination on strategic communication, to interactions with researchers and technologists that can sometimes play key roles in the margins.

Within cyberspace, governments only make up one stakeholder group and it bears reminding that the private sector owns and operates most of the digital and physical assets in any conceivable form, whereas civil society is largely responsible for coding and running the most basic Internet functions that define the parameters of cyberspace. Given the dominant role of these non-state actors, there is a need to move beyond classic like-minded groups of states and consider the role and contribution that civil society and private actors can make. The WoN approach is aimed at facilitating successful *cooperation* between these stakeholders. This includes the exploration of modes of engagement with private companies directly (and not through their "host" nation) that form a crucial part of any deterrence and resilience strategy. Information sharing, including the crucial issue of attribution and intelligence, and interoperability, is a key aspect of all these interactions.

Historically, from the perspective of SMPs the question of how to engage in a WoN deterrence posture is not new. The 'total defense' concept of countries like Switzerland and Austria but also Singapore has always depended on the leveraging of all national assets in times of war.

[189] Stephan de Spiegeleire and Tim Sweijs, "Volatility and Friction in the Age of Disintermediation," *The Hague Centre for Strategic Studies,* (2017); Ernst Hirsch Ballin, Huub Dijstelbloem, Peter de Goede (Ed.), "Security in an Interconnected World. A Strategic Vision for Defence Policy," *The Netherlands Scientific Council for Government Policy* (2020).

For those that were not part of any multi-state security alliance, it was considered the best way to be able to deter larger foes as it enabled the entire society. Even with nations that have decades of experience in leveraging it, the leap to start applying it to cyber is often difficult. For nations without this historic experience, the entire concept can be daunting. While the importance of non-state actors, and the need to cooperate with them, is becoming more widely recognized, one of the biggest challenges to the adoption of a WoN approach is understanding how to apply it. On this matter, two main interpretations have developed. The first one draws insight from the notion of total defense and assumes that the non-state sector is "a government capability in reserve," while the second recognizes it as a "fully capable actor in its own right."[190] As noted in Chapter 3, Russia and China largely consider non-state actors to be subservient to the state. The Chinese Science of Military Strategy describes its cooperation as "in peacetime, use civilians to hide the military; in wartime, the military and the people, hands joined, attack together."[191] In the case of Singapore, the non-state sector is seen as a government capability in reserve (i.e. civil defense). In most liberal democracies, they are seen a fully capable actor in their own right that needs to be convinced to support the government.

Liberal democracies' success mainly depends on their ability to convince non-state actors to voluntarily cooperate.

Based on Joseph Nye's faces of power, three cooperation methods can be identified: coercion, cooption and conviction. Depending on the interpretation given to the role of non-state actors within a nation, the methods to enforce the WoN approach will vary. In case of the government capability in reserve, cooperation will be sought mainly through legislation (coercion) or commercial contracts (cooption), while liberal democracies' success mainly depends on their ability to convince non-state actors to voluntarily cooperate.[192]

First, all governments exert some degree of *coercion* against their citizens and companies through legal instruments. Consider for example European and national legislation that enforces a duty of care and duty to report to operators of critical infrastructure. But regulation can take on stronger forms. Like in Russia, where the SORM legislation requires all Internet Service Providers to install equipment that allows the intelligence services to directly monitor all domestic Internet traffic.[193] Coercion also takes place covertly outside of the legal parameters. In Russia's WoN approach that resembles the Soviet-era notion of 'total defense', the Kremlin instrumentalizes 'hacker patriots', organized-crime groups, businesses, government-organized non-governmental organizations, the media and other actors in the deployment of various 'active measures' in cyberspace.

Second, *cooption* uses positive inducements rather than implied punishment. It includes quasi-volunteer military programs to induce cooperation from non-state specialists. In Israel, a strong private cybersecurity industry cooperates closely with the military and academia "to all three sectors' profit", providing "a potentially powerful mixture of private contracting and a 'whole of nation' approach to cybersecurity and cyber operations."[194] Estonia's paramilitary cyber reserve provides reinforcement for regular military cyber forces in an emergency, as well as competitions organized by the government or military to attract and reward non-state hackers, which are widely used by most mature cyber nations. Such measures are particularly useful given the inability of most nations to maintain all potentially required technical skills in

[190] Alexander Klimburg, "The Whole of Nation in Cyberpower," *Georgetown Journal of International Affairs* (2011), 173.

[191] China Aerospace Studies Institute, "In Their Own Words: Foreign Military Thought. Science of Military Strategy (2013)," Air University (2013).

[192] Klimburg, "The Whole of Nation in Cyberpower," 173

[193] Andrei Soldatov and Irina Borogan, "Inside the Red Web: Russia's back door onto the internet – extract," *The Guardian,* (September 8, 2015)

[194] John Reed, "Unit 8200: Israel's cyber spy agency"; Boeke and Broeders, "The Demilitarisation of Cyber Conflict", 82

their organization at all times. Beyond this operational benefit, it also fulfils a larger strategic and political gain. In China, such programs have been historically popular to attract young programmers and socialize them to the role of the military through the National Defense Reserve Forces, a thirty-year-old military program that includes most computer science students at state universities. Another example is the Chinese 50 Cent Army of Internet users hired by government authorities to act as cheerleaders on online forums – the absolute backbone of the Chinese online civil society. They manipulate online public opinion in favor of the government – proven to be an effective tool for cooption and preventing potential subversives from turning against the state – the CCP's absolute nightmare scenario.

Third, *conviction* is the ability to convince non-state actors to cooperate. In liberal democracies coercion and cooption are only partly effective towards industry due to their limited scope and costs, and are even less likely to work towards civil society. The latter operates independently from government and is often not financially motivated. They also make up the largest and most important part of the wider cyberspace ecosystem, including the technical community, open-source developers, or volunteer 'white hat' and 'gray hat' hackers. Crucially for deterrence purposes, researchers and cybersecurity companies often engage in technical analysis of attacks and attribution. They are often better positioned to isolate and call out foreign threat actors, stigmatize particular transgressions, and mobilize support to impose costs on violators. By publicly delivering 'plausible attribution', these actors can help to deter adversarial state proxies by eliminating the plausible deniability benefit for the government. This also serves as a steppingstone towards more coercive responses against attackers. The first time the EU activated its Cyber Diplomacy Toolbox as a sanction mechanism, the list of entities had already been identified by cybersecurity company CrowdStrike, effectively allowing the EU to sanction without having to do the attribution itself. To this end, "the best defense against non-state hackers who work on behalf of foreign governments may therefore be truly independent and credible non-state researchers, rather than some type of hacker militia."[195] The ability of governments to convince non-state actors of their sensibility and actions depends on its soft power tools and its ability to build a relationship of trust and transparency.[196]

The first time the EU activated its Cyber Diplomacy Toolbox as a sanction mechanism, the list of entities had already been identified by cybersecurity company CrowdStrike, effectively allowing the EU to sanction without having to do the attribution itself.

5.5. **Whole of System, Union and Alliance**

Besides national engagement, there is an obvious need to involve international stakeholders. The Whole-of-System (WoS) approach places the operational center of gravity in an international – rather than national – setting and includes international organizations such as the UN, EU, and NATO, but also multinational companies, such as large Internet platforms or service providers operating across borders, as well as civil society organizations that include the technical community. Hence this approach depends on *collaboration* with a wide range of partners at the international or regional level, whether it is through binding treaties, norms, or non-governmental agreements between non-state actors.[197] The need to develop this approach was especially encouraged by international non-state actors playing a fundamental role in cybersecurity and who tried to increase cooperation with governmental actors while maintaining their independent status. As with the WoN approach, WoS groups traditionally have had a focus on Internet governance through organizations such as ICANN and the IETF.

[195] Klimburg, "The Whole of Nation in Cyberpower", 177

[196] Klimburg, "National Cyber Security Framework Manual," 102

[197] Ibid, 99

They have also focused on countering cybercrime, critical infrastructure protection and crisis management, through cooperation with Internet service providers and the wider non-governmental incident response community (or the firefighters of the Internet). The value of non-state actors in cyberspace can further be seen in scientific and industry working groups, where the role of the government is limited, and efforts to legislate would be unsuccessful.[198] For instance, efforts to tackle the highly advanced and still unexplained Conficker worm in 2008-2010 were led by non-state actors (the Conficker Working Group), with the government largely reduced to watching from the sidelines.[199] The adoption of a WoS approach ensures that operational cyber realities are accounted for, and all stakeholders in cyberspace are considered.

Since the government cannot coerce or compel collaboration (at least in the West), such operational collaboration depends on non-state actors' willing participation.

Within this approach, governments must therefore incorporate the private sector, cybersecurity providers, cloud service providers, telecommunication companies, international organizations, non-profits, civil society, and critical infrastructure owners and operators. As noted by Michael Daniel, "the level of coordination and organization required for effective cyber deterrence policies is much higher than in traditional deterrence efforts. Getting all those divergent actors aligned with respect to goals and activities requires more time, effort, and energy than do traditional deterrence initiatives."[200] Since the government cannot coerce or compel collaboration (at least in the West), such operational collaboration depends on non-state actors' willing participation. To make sure that deterrence does not fall short, a concept of operational collaboration between these stakeholder groups in the form of a WoS approach needs to be put into practice.

It will be important to see how small and medium-sized nations can leverage their national assets at the NATO and EU level. Within the Whole- of Union approach, the integration of sanctions through the EU Cyber Diplomacy Toolbox is the most obvious punishment measure available. Its use is restrained because it requires unanimity from all member states. Further to this, the Mutual Defense Clause (article 42(7) of the TEU) and the Solidarity Clause (article 222 of the TFEU) can offer relief or support to member states that are the victim of a cyber attack. The former is the equivalent to NATO's Article 5 on collective defense and is therefore limited to "armed aggression" or use of force.[201] The Solidarity Clause goes further by creating an obligation on all member states to act jointly and to assist one another in the event of disasters and crises which exceed their individual response capacities. To this end, Cyber Rapid Response Teams (CRRTs), made up of participating member state experts, allow member states to help each other to ensure a higher level of cyber resilience and collective response to cyber incidents. "They can be used to assist other member states, EU Institutions, CSDP operations as well as partners" and are "also able to assist with training, vulnerability assessments and other requested support."[202]

Within the Whole of Alliance posture, it is important to note that the mandate may not only be national: a military cyber organization may receive a mandate to support that nation's allies (e.g. within NATO) in an extension to its common security task. Apart from cyber defense (preparation, response and recovery), this may also include pre-emptive strike capabilities

[198] Ibid.

[199] William Jackson, "Is government the odd man out in cyber defense," *Defense Systems,* (January 31, 2011) https://defensesystems.com/cyber/2011/01/is-government-the-odd-man-out-in-cyber-defense/193155/

[200] Daniel, "Expanding Cyber Deterrence,"6

[201] The European Union, *Consolidated Version of the Treaty on European Union,* C326/15 (October 26, 2012), Art 42.7

[202] PESCO projects "Cyber Rapid Response Teams and Mutual Assistance in Cyber Security (CRRT)" (November 15, 2018).

against a clear and present threat, counter-attack (response), or even an offensive capability mandate. Within the NATO context, cyber capabilities can be integrated as part of Article 5 (collective defense), Article 4 (defensive assistance) and the move towards Sovereign Cyber Effects Provided Voluntarily by Allies (SCEPVA) which are coordinated by the Cyberspace Operations Centre (CyOC). While the Alliance has a purely defensive mandate, offensive cyberspace operations (OCO) can be offered by the member states through the SCEPVA mechanism in accordance with the principles agreed to by NATO.[203] These OCOs may be executed as standalone operations or in conjunction with other operations. However, it needs to be emphasized that SCEPVA, and NATO overall, is more likely to concentrate on a subsection of 'cyber support operations' – namely, battlefield cyber, also known as CEMA. This is because these capabilities are considered very expensive and not widely deployed in the field, therefore increasing the gap between conventional capabilities of partners further. The second reason is that, unlike CEMA, 'strategic strike' cyber is very intelligence sensitive, and while NATO remains an organization of trust, many governments hesitate to share knowledge of their capabilities across all 30 members equally (although such mechanisms already exist between members and are likely to continue, also with outside partners).

It needs to be emphasized that SCEPVA, and NATO overall, is more likely to concentrate on a subsection of 'cyber support operations' – namely, battlefield cyber, also known as CEMA.

The fact that more NATO members now have a credible punishment capability adds to the overall credibility of the Alliance. Yet, other challenges than intelligence sharing persist. First, Allied operations require tight coordination of cyber fires, and the subsequent equities between different nations, intelligence and attack capabilities, which need to be deconflicted at speed. Also, outside the context of allied operations, a country will have to communicate with the alliance before deployment. This process is widely accepted and developed between allied intelligence agencies, but less so when it comes to offensive cyber operations. This is far more complex than coordinating a nuclear SIOP. A common cyber operational picture is much wider and more difficult to obtain even if it were just at the national level. Second, top-tier cyber Allies may consider different kind of cyberattacks that potentially devalue or even invalidate efforts of smaller members – for example, insider or close access operations or side-channel attacks versus more conventional Internet or IP-based operations, respectively. Third, when engagement escalates and comes close to the threshold of conflict or war, one member of the alliance may push the others into a wider cyber conflict. Fourth, in case of a war in which we can expect part of the communication lines between allies and with adversaries to be degraded or damaged, the question of war termination becomes more complicated. Finally, countries involved in alliances may face a strategic dilemma between committing to their cyber capabilities to impose unacceptable costs against the adversary or building up the capabilities of the alliance. This is especially pertinent for smaller and medium sized nations that face the larger challenge of increasing their leverage vis-à-vis the alliance.

5.6. Organizational Requirements and Considerations

Having considered the need and ways in which SMPs can work towards a Whole of Government approach to cyber deterrence, and how it affects their position in an alliance, an assessment follows of the organizational considerations that such a posture requires. This ranges from attribution, interoperability and common definitions, narrative control and defined strategic objectives, the intelligence capabilities and the strategic cyber weapon

[203] NATO, *Allied Joint Publication-3.20: Allied Joint Doctrine for Cyberspace Operations*, Edition A, Version 1 (January 2020).

development itself, the management of second and third-order effects and equities tradeoffs, and coordination and deconfliction with international partners

First, deconflicting terminology. While a common language is always preferred among allies, for example by adopting NATO definitions, this comes with costs as well. Having a common understanding of cyber terminology is the first step towards interoperability among partners at the strategic level. But typologies and definitions are perhaps one of the most elusive components of cyber operations due to the fast-changing nature of the domain. An overemphasis on definitions is therefore not helpful either. Instead, clarity of communication and concepts is extremely important, as is a focus on intended or experienced effects, rather than becoming bogged down in definitional quarrels by excessively emphasizing technical or legal aspects. It is often much easier and more useful for policymakers to develop useful descriptions outlining general concepts, rather than fixating on tight definitions. At the same time, the significant differences in knowledge on cyber capabilities, both across allies but also within governments, makes common terminology sometimes an unsurmountable obstacle. This is accentuated by the fact that that the United States, whose experience in technical cyber operations is unmatched, has sometimes defined cyber operations not only at odds with the knowledge of allies, but also their use and employment by adversaries. Put otherwise, some cyber operations experienced by Western nations may not be easily explained by using current high-level doctrine, and may require a more nuanced explanation to decision makers.

Process to non-state actors, and that is to increase its credibility – at the same time, it is unlikely that extreme measures (counter- and cyber attack operations) can be argued with private sector attribution alone.

Second, there is the attribution test that not only requires small to medium-sized states to be able to technically attribute bad behavior to a state actor with a sufficient level of confidence, but also to convince others that they are in fact able to do so. The technical component is only part of this test – one which the US and several of its partners are convinced has been solved. Communicating this ability can happen in several ways. States can, for example, openly disclose how their operational intelligence collection and analysis functions work. Hardly a single government has voluntarily disclosed this kind of information, and there is very likely an extraordinary knowledge gap between public (and even the 'knowledgably public') and the intelligence reality. The attempt to obfuscate all methods and means was likely an error and has made governments pursue another option. In short, the 'trust us' message. This, however, only works in favorable political circumstances, and most leading liberal democracies have seen significant challenges to their credibility emerge over recent years.

From the outside, it is difficult to accurately evaluate how credible a state's ability to attribute is. And while there are no international standards of proof for most of the retaliatory actions in cyberspace, countries still have a strong incentive to not make spurious allegations, lest they lose credibility. Overall, increasingly more states are engaged in public attribution, albeit with limited disclosure of the technical details to preserve intelligence assets and techniques. Some offer insights into their process while others show off their capability through (counter-) intelligence operations.[204] There can also be an advantage to outsourcing the attribution process to non-state actors, and that is to increase its credibility – at the same time, it is unlikely that extreme measures (counter- and cyber attack operations) can be argued with

[204] When the UK condemned Russia's GRU over a Georgia cyber-attack, a framework used by the UK government for all source assessments, including the probability yardstick, was published as well, albeit no longer publicly available. UK Government, "UK condemns Russia's GRU over Georgia cyber-attacks", (February 2020), https://www.gov.uk/government/news/uk-condemns-russias-gru-over-georgia-cyber-attacks. It was also reported that the Netherlands intelligence service AIVD conducted an intelligence operation against Russian hacking Group Cozy Bear, which is associated with the GRU, and watched Russian hackers launch an offensive cyber operation against the US State Department. Rick Noack, "The Dutch were a secret U.S. ally in war against Russian hackers, local media reveal", The Washington Post (January 2018). https://www.washingtonpost.com/news/worldviews/wp/2018/01/26/dutch-media-reveal-country-to-be-secret-u-s-ally-in-war-against-russian-hackers/

private sector attribution alone. Also, while it is possible to identify with varying degrees of certainty who was behind the attack, it might not be possible to do it in a politically meaningful timeframe, especially in situations of conflict. This is particularly pertinent when taking retaliatory actions through cyberspace that cannot be undone once unleashed. To the contrary, sanctions or indictments can be rolled back in case of misattribution. The time requirements to achieve attribution at a reasonable confidence threshold are considerable create temporal breaks for the pressure of a crisis situation to diffuse and for decision-makers to evaluate alternative courses of action.[205] This process includes the willingness to share potentially sensitive intelligence information with allies to justify any response, and following this, the evaluation of the shared intelligence assessment by the allies.

Third, narrative control and defined strategic end goals are the leading factors when responding through cyberspace, ranging from cyberoperations to information warfare. This means that the political and strategic objectives– not the tactical and operational objectives – are front and center. For example, if the narrative when responding to an attack is one of injured innocence, overt and possibly even covert offensive responses can be a setback. Alternatively, if the narrative is 'resolute defense' then disproportionate retaliation early on is important. The centrality of narrative avoids the bottom-up problem of placing the tactical before the political, which has been an essential part of US path dependency in cyber operations.[206] This refers back to the natural trend favoring the technical operators in a highly complex and esoteric field as cyber and a "bottom-up culture of putting technical feasibility before political desirability, which is hardwired into the NSA and US Cyber at large."[207] While a bottom-up process is often considered a positive development because it is based on a sound technical basis, in some cases it can lead to strategic decisions being taken within a very narrow (and low level) strategic framework. In other words: "Changes to facilitate a particular task at this level can, however, greatly impact the core values of a nation. This can occur without the strategic or political level being fully cognizant of what is occurring."[208] Making the political and strategic goals key is not only a way to avoid path dependency, but it also helps in establishing redlines and a wider assessment of the equities that, in contrast to tactical assessments, may lead to a different response because of the broader strategic or geopolitical environment.

Making the political and strategic goals key is not only a way to avoid path dependency, but it also helps in establishing redlines and a wider assessment of the equities that, in contrast to tactical assessments, may lead to a different response because of the broader strategic or geopolitical environment.

Narrative control starts before an operation is even considered. One of the main lessons learned from the rollout of the US defend forward strategy and persistent engagement doctrines is the need for consistent strategic communication and signaling of a strategy inside all branches of government, from the political to the tactical level, and to outsiders, from allies to opponents: "One of the crucial deficits that emerged from the [Solarium] commission's research is that there is confusion among multiple audiences—including within the US government—and inconsistencies in official documents about strategic approach definitions and end states."[209] For example, "active disruption" and "persistent denial" are used in the 2019 National Defense Authorization Act, whereas USCYBERCOM's vision uses "persistent engagement", while the DoD Cyber Strategy uses "defend forward", which in turn is not even mentioned in the National Cyber Strategy from the same year. Their strategic objectives and nature also differ widely, from maintaining American superiority and military advantages to

[205] Erica D. Borghard and Shawn W. Lonergan, "Cyber Operations as Imperfect Tools of Escalation," *Strategic Studies Quarterly* (2019). https://www.airuniversity.af.edu/Portals/10/SSQ/documents/Volume-13_Issue-3/Borghard.pdf

[206] Klimburg, *The Darkening Web*, 200.

[207] Ibid, 149-150

[208] Klimburg, "National Cyber Security Framework Manual", 121

[209] Lonergan and Montgomery, "Defend Forward as a Whole of Nation Effort,"

protecting critical infrastructure and international stability. Within the Pentagon it is portrayed as a defensive strategy, but at the same time John Bolton, the former National Security Advisor, described it as: “we’re going to see more aggressive offense from the US side ... like retaliation.”[210] Strategic communication and signaling efforts should therefore be synchronized across all branches of government, rather than inconsistently depending on the military. The Ministry of Foreign Affairs plays a key part in these efforts, either through its public diplomacy or private channels, and must be integrated in any deterrence effort or military concept for that matter.

One of the most pertinent intelligence reforms is to open the black box that is cyber conflict to decision-makers everywhere – first and foremost political, but also key national leaders in business and civil society that need to be brought on board.

Fourth, intelligence analysis and common operational pictures require a much stronger level of investment than traditional security concerns. Cyber, more than any other military tool or even tool of statecraft, is intelligence dependent. Most SMPs (even as a ‘medium’, Tier 4-6 level cyber power, explained in Annex III) may have little challenges in actually creating offensive cyber teams (see point 5), but resourcing the intelligence needed might be more difficult. Intelligence is required to understand the threat of actors and their motivations; their tools techniques and procedures; their order of battle; and ‘armament’ and supporting infrastructure and supply chain. Specific targets and target sets – from individual adversarial critical assets to interlocking infrastructures – need to be subject to close-quarter reconnaissance, literally the mapping of the network. Equally, however, the adversarial nation state needs to be fully graphed as an interlocking political system, for instance using the PMESSI system described above, so that the actual target selection makes sense. Part and parcel of this is also the international dimension, where all actions will have repercussions, depending on the visibility and severity. All of these tasks need to be supported by one’s own collection and partner collection, and itself synthesized from a very wide range of different intelligence sources depending again on a variety of relationships to be effective. Creating a cyber common operational picture is sometimes incorrectly viewed as the end goal – however, this misstates both the challenges in developing the picture, as well as the utility. The challenges are immense, and the utility is not just to target cyber capabilities better but rather to communicate the extent of the larger conflict arena to political decision-makers. One of the most pertinent intelligence reforms is to open the black box that is cyber conflict to decision-makers everywhere – first and foremost political, but also key national leaders in business and civil society that need to be brought on board. This would constitute a major contribution to the strategic culture. Such reforms should not only extend to the executive branch, but also at its interaction with the legislative branch. This includes an institutional basis that allows for confidential briefings to members of parliament with the necessary security clearance, much like the US Senate or House Committees. Capacity building within government – informing leading civil servants of the challenge ahead – is a task seldom completed. A common cyber operational picture is therefore much more a ‘picture’ than anything else.

Fifth, organizing offensive cyber means. As mentioned previously, the US Defense Science Board report clearly shows that cyber capabilities can start very rudimentary – a lone hacker, appropriately skilled and equipped – and can extend to dozens of teams (each with up to dozens of individuals) and hundreds of millions of euros worth of resources.[211] Put differently, cyber capabilities and their cost increase on a logarithmic scale. While the very top of the pyramid (the sole Tier 6 cyber power that is the US) spends billions of euros on cyber every year, much less well-equipped Tier 3-4 actors can still have a deterrent effect with a fraction of that investment. A cyber campaign needs to have intelligence on the target, different tools

[210] “Transcript: White House Press briefing on national cyber strategy.” *Grabienews,* (September 20, 2018)

[211] Depart of Defense, Defense Science Board, “Task Force Report: Resilient Military Systems and the Advanced Cyber Threat”, 21-25

(including zero-days if needed) and infrastructure for initial intrusion, a separate range of tools for moving within the target, and finally, one or more 'weapons' that need to be deployed without 'inadvertent release' occurring. The cyber weapon itself can be an off-the-shelf piece of malware (like "Wiper"), or it can be a highly bespoke piece of code produced in a dedicated cyber foundry, tested on fake systems, and modeled in a test range. As a general rule of thumb, the level of preparation for a cyber operation is directly proportional to the level of distinction and limitation of second-order effects. There are few if any off-the-shelf cyber bombs, and any accurate use requires a certain level of advanced planning and intelligence gathering. Libicki estimates that the time and resource requirements for intelligence gathering on a target exceed those needed to conduct an operation potentially by a ratio of 100 to 1.[212] Only the most indiscriminate operations can be expected to be launched with minimum preparation, although at times even these operations might still be desirable and can still have a deterrence effect. Doing cyber legally, in full accordance with international law and laws of armed conflict, is simply much more expensive and time consuming, but also leads to a higher class of offensive cyber capability.

Sixth, targeting and second and third-order effects. Compared to conventional weapons, cyber operations can target more accurately and with less physical collateral damage. At the same time, the propagation of second and third-order effects is so extensive that only biological warfare can claim to be even less controllable. After all, cyber effects can occur outside the area of operation, in domains other than cyberspace, and can undermine the security of other (neutral, allied or civilian) actors that rely on the same system – NotPetya being an infamous example – or the Internet overall. The interconnectivity and interdependence of military, civilian, and private networks and systems increases the risk of unintended effects to non-combatants, non-military objects, or the wider information environment across the world. This make it much harder to determine the likelihood and risks of unintended effects in cyberspace compared to the collateral damage assessments of conventional weapons where military planners may be able to estimate, for example, the probability of an incoming weapon missing its military target and hitting a nearby civilian facility. All effects – intended and unintended – are only as good as the available intelligence on the target. Minimizing unintended effects in cyberspace requires not only detailed intelligence about the target system, but also its relation to other systems. Furthermore, a 'cyber weapon' can sometimes also be reverse engineered by the victim, often leading the attacker to tailor the malware in such a way that it wipes itself from the networks once it is revealed. This makes it even harder to get a full understanding of the offensive tool before it is used, which further increases the likelihood of unanticipated effects. These wide-ranging effects narrow the distance between the tactical and strategic levels and must be factored into the planning of cyber operations. The expected military benefit that a cyber operation may produce must be weighed against the estimated amount of collateral damage or disruption and be considered proportional. This balancing act is largely informed by the circumstances. After the operation, the collateral effects also need to be managed, often through the involvement of allied states and industry partners, so patches against the used vulnerability are implemented by allies, requiring a Whole of System information-sharing mechanism.

Beyond these operational considerations, states should also limit long-term (normative) second and third-order effects that undermine international law, international humanitarian law, norm processes within the UN and elsewhere, and Internet governance. As argued previously and elsewhere, offensive cyber operations may introduce new norms or precedents

[212] Libicki, "Cyberdeterrence and cyberwar", 155

that may harm the long-term strategic interests of liberal democracies.[213] For example, the public-facing component of the US defend forward strategy and persistent engagement doctrine oriented itself around the overt imposition of costs, directly compromising Russian troll factories and using coercive signaling via pre-deployment in its electrical grids. It thereby conveyed a public message implying that it is now acceptable to hack what you consider 'fake news' thereby encouraging disputes about 'bad content'. Ultimately, this may lead to the very thing the doctrine was intended to alleviate: the weaponization of information. Furthermore, by openly communicating about their pre-deployment (rather than being caught in the act), it designated critical infrastructure as a viable vector of coercive signaling – that the range of acceptable cyber targets had expanded to include critical infrastructure, up to the point of threatening 'mutually assured disruption.' Without recognition of these long-term effects upon the wider international legal order in cyberspace, unintended consequences may undermine the very goals states wish to achieve and set dangerous precedents that will likely inform future calculations of other actor's behavior in cyberspace, rendering the broader cyberspace environment more uncertain, hostile, and complex.

Seventh, the equities tradeoff presents several dilemmas. The collateral damage assessment for unintended effects that can undermine the cybersecurity of third parties is one of them. Related to this are the political or legal ramifications that may result from cyberattack operations conducted in hot pursuit of a target positioned in neutral or gray space. Perhaps the most well-known dilemma is between cyberattack and ISR operations, whereby the latter occurs covertly, and the former's destructive effects means that previously unknown vulnerabilities likely become apparent to the target. Or, as Lonergan explains, the strategic worth of intelligence – a precursor to support offensive operations – "demands that governments conduct intelligence gain/loss calculations when evaluating the potential upside of conducting offensive operations that may jeopardize cyber intelligence assets."[214] This not only means that these intelligence assets can no longer be used by attackers, but when an offensive operation shuts down a system, it can also halt an ongoing intelligence operation from an ally on the same system. Similarly, larger partners may execute operations that may hinder operations of smaller partners – such things are not always coordinated, and they may contradict each other's national interests. This requires communication and coordination with partners and allies, and a decision on to what extent they are informed or asked for permission before an operation. The same equity question can arise within a single government as well, where one branch might be more interested in maintaining presence and intelligence gathering, while another is more interested in 'finishing' a target.

Worth of intelligence – a precursor to support offensive operations – "demands that governments conduct intelligence gain/loss calculations when evaluating the potential upside of conducting offensive operations that may jeopardize cyber intelligence assets."

Another dilemma is known as NOBUS in the US – a term used by the NSA to determine the likelihood that an adversary can exploit the same vulnerability or if nobody but the US (NOBUS) can exploit it. This operational dilemma can also lead to another one: the 'use it or lose it' quality of a vulnerability in which there is an apparent limited window of opportunity for its exploitation, possibly making it more likely for a state to use any offensive capabilities it may possess in a given moment out of the fear that they will not be available for future use. As noted by Lonergan, "while these incentives may exist when stakes are high, or for decision-makers with certain risk profiles, the reverse is also true: using a capability nearly guarantees that it won't be available for future use."[215] These assessments increase or reduce the probability of escalation through cyberspace – using an offensive cyberattack operation as

[213] Alexander Klimburg, "Mixed Signals: A Flawed Approach to Cyber Deterrence"; Louk Faesen et al., "From Blurred Lines to Red Lines. How Countermeasures and Norms Shape Hybrid Conflict"

[214] Borghard and Lonergan, "Cyber Operations as Imperfect Tools of Escalation,".137

[215] Ibid.,135

means of deterrence by punishment – contingent on how decision-makers rank the related utilities in their equities process.

Throughout all these dilemmas, a state needs to recognize considerable differences across its allies and partners in terms of their own offensive and defensive capabilities and their political willingness to conduct attributable versus unattributable cyberattack operations for the sake of interoperability. When operating in gray or blue space, the state needs to be aware of entanglement and differences in the application of international law to cyberspace (e.g. in sovereignty) and their willingness to allow foreign – albeit allied – operations in their own network, their preference for notification prior to, during or following such a foreign operation. The Ministry of Foreign Affairs plays a crucial role in these considerations and in mobilizing allied support for the overall strategic objects and political deconfliction, taking into account the diverse points of view and capabilities of their partners and opponents alike.

Eighth, interoperability, summarized as the ability to be informed, to support, and to execute, is a crucial precondition and challenge for cyber operations that are part of a multi-agency or coalition operation. These three abilities reflect the level of ambition, starting with the ability to be informed by allies and ending with the ability to execute offensive cyber operations. At least, interoperability requires an understanding of the resources, capabilities, goals, strategies, doctrines and ideological context of another state. In the NATO context, interoperability often refers to the ability to act jointly in a coherent, effective and efficient manner to achieve tactical, operational, and strategic objectives. At the strategic level, this is facilitated by common terminology, strategies and doctrines, and at the operational and tactical levels by aligning TTPs. Finally, technological interoperability provides compatibility between actors' ICT in the areas of information assurance, command and control, sensors, and firepower. Not only is the underlying technology often an obstacle to interoperability, but different information security policies can prevent or limit the exchange of information. One tool that contributes to such interoperability is standardization (or a standard-based approach), which allows international units to cooperate through aligned procedures, concepts, equipment, and information technology.

Given the invisibility and opacity of cyber means, one of the most challenging components of cyber deterrence is demonstrating your offensive cyber capability and the will to use them in such a way that it deters potential aggressors.

Ninth, communicating offensive cyber capability. Given the invisibility and opacity of cyber means, one of the most challenging components of cyber deterrence is demonstrating your offensive cyber capability and the will to use them in such a way that it deters potential aggressors, rather than incentivizing them to develop and use their own cyber capabilities as a response. First, opponents need to be aware of your deterrence capability. After all, Dr. Strangelove teaches us you can't deter with secret weapons. Compared to conventional weaponry, there are considerably fewer options available to nations in the cyber context. Often, the mere existence of offensive cyber capabilities, and their effects, cannot be known with confidence until after they have been used and their effects become apparent to the target or a wider public. States are therefore left guessing as to the overall capability of another state (albeit at widely varying degrees of detail) without, for the most part, being able to determine the exact order of battle, table of equipment, tactics, techniques, procedures or other basic information – unless the intelligence assessment is very complete. In turn, some of the conventional demonstration methods, such as experiments or examinations, have diminished returns, while others have become more important. To this end, two demonstration tools stand out: exercises and employment. First, one way to communicate capabilities is through cyber exercises, reported on in the media. This includes national and allied exercises, such as the NATO's Cyber Coalition 21, Crossed Swords and Locked Shield (although such exercises often have a defensive focus). It requires strong and coordinated strategic communication, ranging from government disclosures, leaks, or (confidential) briefings. The US has often revealed capabilities through (mis)reporting in the wider media. This approach, however, not only requires very tight message control, but it also needs to follow a very fine line between propaganda and public diplomacy. Most importantly, it must be wary of instrumentalizing the free press as well as civil society to spread untruths – for instance on one's relative strengths or weaknesses – as this would clearly cross a fine line that liberal democracies should instead be actively trying to defend.

Second, based on the assumed premise that cyber operations hardly, if ever, escalate into kinetic conflict, some nations and experts argue in favor of employing offensive tools as a form of deterrence by punishment. This can include demonstration strikes and other means of signaling intent and capabilities. This would signal not only the power of offensive cyber capabilities and the will to use them, but also the ability to find vulnerabilities in the other's systems. The attacker can often decide to what extent they want an attack to be visible – and even decide to what extent they want it to be attributable to them. For instance, in many cases the attacker may wish to give the defender a 'political way out' by keeping the attacks and the damage caused out of the public eye – but at the same time they may want to make sure the defender is aware of who the attacker is. Sometimes such *discrete punishment*, not publicly visible, but attributable to you, may be the best option to signal resolve without pushing the adversary into a public relations corner. *Anonymous punishment* – non-attributable to you specifically but maybe to your allies in general – can be a particularly useful strategy to a small or medium-sized state that is part of a larger alliance – and detrimental to a nation without them. Employing offensive tools can, however, risk incentivizing the other side to respond in kind and contribute to a tit-for-tat escalation. Signals through cyberspace can be open to multiple interpretations, they can be lost, or only found out after a long delay. In what Healey describes as the Cartwright Conjecture, strong capabilities, and the will to actually use them in conflict seemed to have little deterrent value. He recalls the US-Iranian example that resulted in worse outcomes: "The original attacks on Iran of Stuxnet and Wiper (Flame) not only seemingly failed to deter the Iranians, but rather caused them to counterattack. Worse, the proportional responses against US banks and Shamoon were then used as "wake up calls," then used as justification by America's cyber warriors to advocate for more cyber capabilities. [...] Instability and escalation seem to be the theme, rather than deterrence or restraint."[216] In a – yet untested – alternative route, threats and attacks are specifically intended for deterrent purposes or clearly defined redlines from a state. This means that a capability would not just be a demonstration of force, but clearly linked to behavioral benchmarks. Deterrence requires some clarity on where the redlines are and how willing such a state is to carry out its threat and by what means. Absent such clarity, brandishing may have an effect opposite from the one intended.

Deterrence requires some clarity on where the redlines are and how willing such a state is to carry out its threat and by what means. Absent such clarity, brandishing may have an effect opposite from the one intended.

5.7. **Main Takeaways**

Cyber deterrence is too multifaceted an issue to be dealt with by just the military. To encompass all the various stakeholders, three different approaches to cyber deterrence are required: Whole of Government to facilitate *coordination* among government agencies, Whole of Nation to facilitate *cooperation* between national state and non-state actors, their civil society and industry partners. Whole of System to facilitate *collaboration* with a wide range of

[216] Jason Healey, "The Cartwright Conjecture: The Deterrent Value and Escalatory Risk of Fearsome Cyber Capabilities," (June 28, 2016) https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2836206

state and non-state partners at the international or regional level, most notably within NATO (Whole of Alliance) and EU (Whole of Union) context.

Because of their limited resources, SMPs would benefit greatly from leveraging their wider national non-state assets. Liberal democracies are more limited in coercing or coopting cooperation from non-state actors (industry to a certain extent but especially civil society). Instead, success will largely depend on their ability to convince these stakeholders.

While Whole of Government is the predominant approach in cyber deterrence, the two other approaches can more easily help governments attain this ideal state. This includes high-level changes to governmental national security policy, like the introduction of a National Security Council type structure where appropriate.

Because of their limited resources, SMPs would benefit greatly from leveraging their wider national non-state assets.

Despite the noted importance of collaboration, there are enduring challenges that often hamper effectively cohering the different stakeholders. To this end, we offer nine organizational recommendations for minimum deterrence ranging from attribution, interoperability and common definitions, clear political objectives and end-game, intelligence capability, mandate and legality, the strategic cyber weapon development itself, coordination with partners and allies, and the management of collateral effects.

1. **Align common terminology or at least clarity of communication and general concepts** with allies and partners. Typologies and definitions are perhaps one of the most elusive components of cyber operations. It is much easier and more useful for policymakers to develop useful descriptions outlining general concepts, rather than fixating on tight definitions This is a crucial first step towards interoperability as SMPs' cyber operations will often operate as part of an alliance structure.
2. **Develop and communicate the ability to attribute in a politically meaningful timeframe**, in which SMPs are able to technically and politically attribute aggressive behavior from a state actor with sufficient confidence levels, and are willing to share intelligence domestically and with allies to legitimize a response.
3. **Define strategic-end goals and prioritize narrative control** over tactical and operational objectives to avoid path-dependency and the bottom-up problem of putting technical feasibility before political desirability. Narrative control starts with the formulation of a strategy, and requires synchronized signaling and strategic communication across all branches of government.
4. **Expand intelligence capabilities** that go beyond the minimum requirements of a common operational picture, which requires support from partner collection and a higher level of resources to beyond the operational vantage point. Intelligence is required not only to understand the tactical and operational side of the target (their threat, motivations, TTPs, order of battle, armament and supporting infrastructure, their networks), but to graph their interlocking political system, for instance by using the PMESSI framework, so that the actual target selection makes sense. Creating a cyber common operational picture is sometimes mistaken as the end goal – however this misstates its utility. This is not just to target cyber capabilities better but also to communicate the extent of the larger conflict arena to political decision-makers. Perhaps one of the most-needed intelligence reforms is to open the black box that is cyber conflict to decision-makers. This should not only be directed at the executive branch, but also the legislative branch, which should, for example, be able to receive confidential briefings on offensive cyber developments, provided they are thoroughly vetted. Within some SMPs such an institutional basis is missing.
5. **Invest in offensive cyber means capabilities**, including the tools and infrastructure for initial intrusion, for moving within the target, and finally for one or more 'weapons' that need

to be deployed without inadvertent release occurring. The costs of cyber capabilities increase on a logarithmic scale, starting at a rudimentary level (indiscriminate, off-the-shelf malware with minimum preparation) going up to the top of the pyramid (discriminate, bespoke malware with lengthy preparation). Many SMPs may lack the resources and intelligence of the top-tiered cyber powers to conduct those high-end operations but can still establish a deterrent effect through less sophisticated means.

6. **Assess and limit second and third-order effects,** which are much harder to determine given the complex and dual-use nature of the cyber domain. They can take place outside the area of operations, in other domains than cyberspace, halt allied operations, and undermine the security of other (neutral, allied or civilian) actors. A detailed intelligence assessment of the target system and its relation to other systems can minimize these operational risks. States also need to take the unintended (normative) effects of cyber operations into consideration that – especially if taken overtly – can establish dangerous precedents that undermine their long-term strategic interests.
7. **Assess and balance dilemmas in the equities process,** such as the second and third order effects, how it impacts other – often intelligence – operations with the same target, how long the vulnerability is likely to be available, and whether others are aware of the same vulnerability, to name just a few. These assessments are contingent on how decision-makers rank the respective utilities in their equities process, while considering the diverse points of view and capabilities of partners and opponents alike, again requiring international coordination.
8. **Strive for interoperability.** SMPs will often operate as part of multi-agency or alliance structure, which require them to synchronize, to be informed, to support and to execute cyber fires with allies. This is facilitated by common terminology, understanding of the resources, capabilities, goals, strategies, doctrines and ideological context, as well technological interoperability across ICT systems and policies.
9. **Communicate your offensive cyber abilities** without causing mixed signals or unintended escalation. One way this can be accomplished is through cyber exercises, reported on in the media. Another includes demonstration strikes as a means of signaling intent and capabilities. This would signal not only the power of offensive cyber capabilities, but also the ability to find vulnerabilities in the other's systems. By clearly linking offensive use to previously establishing redlines, SMPs can tie deterrence efforts to more clearly demarcated thresholds, thereby using their punishment capability in a way that minimized the risk of misinterpretation and inadvertent escalation. Furthermore, in a conflict scenario, sometimes discrete punishment (not publicly visible, but attributable to you) may be the best option to signal resolve while at the same time not pushing the adversary into a public relations corner.

6. Conclusions and recommendations

The technological paradigm shift that is caused by the emergence of cyberspace has created a unique opportunity for advanced SMPs to engage in deterrence through cyberspace. This leads to an unprecedented reality that many such smaller nations do not yet fully grasp and which bears significant challenges to both national security organization and culture.

Why should SMPs engage in deterrence through cyberspace at all? Traditionally, they relied on the protective umbrella of their larger allies, and their strategic mindset was directed at resilience and the fostering of norms of behavior and situations of entanglement to avoid conflict. However, the rise of cyber conflict has flattened what used to be considered a deterrence 'ladder' into a deterrence 'vortex'. This means that all aspects of deterrence have to be exercised simultaneously for any one of them to be successful. Entanglement, norms, resilience and punishment all have to be present in a nation's strategy for any one of them to be a success. The current geopolitical context has also changed since the Cold War. The geopolitical environment has become much more hostile and the range of threats governments face have become all-encompassing, fall into blurred areas, and are no longer largely defined by geography. Alliance structures are more complex and arguably do not guarantee the same level of protection anymore, as interstate competition and conflict deliberate avoids thresholds that would trigger a collective defense.

Against the backdrop of a more hostile geopolitical environment, SMPs have a capacity not only to defend themselves but also to threaten larger adversaries in turn.

Against the backdrop of a more hostile geopolitical environment, SMPs have a capacity not only to defend themselves but also to threaten larger adversaries in turn. Some of these nations may not be aware of their innate capability to engage in strategic retaliation. Whether they like it or not, cyberspace has just changed their deterrence ability, just as it has changed the overall conflict landscape. In fact, other states may actually presume that SMPs with advanced economies already have the offensive cyber capabilities to engage in deterrence, even if the strategy of the latter is to hide those capabilities. SMPs are expected to engage in this capability. Even in states where some thinking has already taken place in this regard gaps remain, not just when it comes to cyber capabilities, but also, and more importantly, to the strategic culture and intelligence reforms that are needed to support this. A posture of minimum deterrence through cyberspace requires the development of concepts and capabilities at the strategic, operational and tactical levels, the legal frameworks, and organizational structures. The analysis in this report offers the following takeaways:

First, deterring *through cyberspace* – is more complicated compared to conventional or nuclear deterrence. In cyberspace, the range of threat actors and targets is much larger and now encompasses both state and non-state actors and the hybrid forms in between. States are not monolithic entities - many different government departments engage in cyber operations leading to a cacophony of action, not only from varying mandates within government but also due to the activities of proxies and other state-affiliated organizations. Not only is the affiliation of the actors unclear but they can be multiple at once. Cyberspace is also much less transparent, perpetrators are harder to identify, signaling one's intentions and capabilities is more complex and prone to misinterpretation and therefore, inadvertent escalation.

And because it remains a new and unique domain, thresholds and redlines remain unclear while a deterrence grammar is being developed. This poses challenges to the development of a robust cyber deterrence posture that small and medium-sized nations will need to tackle.

Second, the nature of deterrence has changed, meaning that promoting norms alone will not be sufficient, especially in a world of continuous low-level conflict. In previous decades, dealing with the lower levels of the deterrence ladder was perhaps sufficient for SMPs, but the rise in overall belligerency and interstate 'below the threshold' conflict as well as the larger diffusion of power to the margins (nonstate actors and smaller nations) means that deterrence has changed as well. The flattening of the deterrence ladder into a vortex means that all four aspects – norms, entanglement, denial, and punishment - must be equally engaged for deterrence to work, as 'that which is deterred' is not only civilizational-ending nuclear exchanges but much more mundane and consistent violations of sovereignty and infringements on national interest. The greatest change here for SMPs is therefore their need to engage in deterrence by punishment. However, the technical challenges in doing so are minor compared to the organizational or cultural challenges.

The sheer number of nations that may be able to compete with and reciprocally threaten a major power could be historically unprecedented.

Third, SMPs now possess something that was historically unavailable to them: a strategic weapon capability. Costs and capabilities for offensive cyber operations increase on a logarithmic scale – while very top cyber powers may well spend billions of euros towards this end, smaller and medium sized nations with advanced economies can easily field an "acceptable" minimum strategic capability for a fraction of these costs. This means that, for the first time since nuclear weapons were developed, these nations can actively wield a kinetic-equivalent deterrence by punishment option. This form of deterrence does not seek to prevent cyberattacks from ever occurring, but is aimed at integrating cyber operations as a punishment capability within an overall deterrence posture that also calls on the other three dimensions of deterrence to work. In this new deterrence world, not only 'acts of war' need to be prevented but also consistent and sustained levels of hostilities below the threshold of armed conflict. The destructive potential of these SMPs may be orders of magnitude less than that of the US or near-peer cyber powers, but some of them still possess a minimum deterrence capability that – much like the small nuclear arsenals of France or the UK– could inflict an unacceptable level of retaliatory punishment on a larger aggressor, no matter their overwhelming technical superiority. At the same time, the ability to retaliate to activities below the threshold of the war is equally important as the ability to prevail if such a war does occur.

Fourth, the objective of a minimum cyber deterrence capability for SMPs is not the ability to win a war, but to inflict unacceptable costs to another actor and dissuade them from engaging in hostile cyber acts, both below and above the threshold of war. This does not mean that such deterrence merely rests on the certainty of inflicting unacceptable damage on an aggressor, but on the potential aggressor's uncertainty of avoiding unacceptable damage. Cyber operations can be a relatively cheap way to achieve this objective, both in "special operations" (covert action) below the threshold of war as well as cyberattack operations in an armed conflict. The sheer number of nations that may be able to compete with and reciprocally threaten a major power could be historically unprecedented.

Fifth, in addition to countervalue and counterforce targets, a minimum cyber deterrence also introduces a third category: *counterpolitical* targets. Traditional adversarial counterforce (military) and countervalue (wider society) targets are more likely to establish overt effects that are noticed by the target. Counterpolitical targets, on the other hand, offer a more covert and less escalatory avenue. They include a narrowly defined target of high political value, such as individual oligarchs or companies that hold high intrinsic and symbolic value within a country, even if the value is completely hidden from public view and not widely shared. Informed by targeting practices of bespoke economic sanction regimes, such an approach accepts the information warfare narrative that is being pushed by some states and which undergirds their overall political objectives and strategy, but refuses its means – namely, subverting the free press and systemic media ecosystem with disinformation and malinformation.

Sixth, the intelligence requirements extend beyond the standard operational approach and include psychological, political, economic and other considerations. To this end, the OSCOPE-PMESII framework can be used. As a general rule of thumb, the level of preparation for a cyber operation is directly proportional to the level of distinction and limitation of second-order effects. Only the most indiscriminate operations can be expected to be launched with minimal preparation, although at times even these operations might still be desirable and can still have a deterrence effect. Doing cyber legally, in full accordance with international law and laws of armed conflict, is simply much more expensive and time consuming, but also leads to a higher class of offensive cyber capability. At the same time, the concentration on a limited number of targets rather than a wider range of counterforce and countervalue targets, conserves intelligence resources for maintaining access. There is, however, an additional intelligence lift. Traditionally, intelligence needs and preparation requirements occur mostly on the operational level to create an operational picture of the target, without much attention to the political, economic or other considerations that can contribute to determining the political value of a target. To fill this gap, the counterinsurgency ASCOPE-PMESII framework can function as a useful tool to determine the wider intelligence needs required for better understanding of the wide operational environment that can contribute to determining the political value of a target. To this end, inflicting unacceptable damage depends on the political objectives and 'nightmares' of the target - some countries may find temporary shutdowns of critical infrastructure unacceptable, while others may consider their own regime security as sacrosanct.

Inflicting unacceptable damage depends on the political objectives and 'nightmares' of the target - some countries may find temporary shutdowns of critical infrastructure unacceptable, while others may consider their own regime security as sacrosanct.

Seventh, given the diversity of the actors involved, a Whole of Government, Whole of Nation and Whole of Society approach is needed for cyber deterrence. While Whole of Government is the predominant approach in cyber deterrence, the other two approaches can help governments attain this ideal state more easily. This includes high level changes to governmental culture and national security policy, like the introduction of a National Security Council type structure where appropriate. Furthermore, because of their limited resources, SMPs would benefit greatly from leveraging their wider national and international non-state assets. Liberal democracies are more limited in coercing or coopting cooperation from non-state actors (industry to a certain extent but especially civil society). Instead, success will largely depend on their ability to convince these stakeholders to work together with government.

Eighth, in order to achieve a cyberattack capability that can function as a minimum cyber deterrent, a number of organizational requirements are suggested. From all these considerations for a cyberattack operation, three components are worth highlighting: the level of preparation, working with ambiguity and unintended effects. The careful preparation of an offensive operation will significantly enhance its success, as the level of preparation is directly proportionate to the protection level of the target and the distinction (limitation) of the preferred effects. But even the best preparation will only go so far in addressing the second distinguishing feature: managing ambiguity and inadvertent effects. This challenge starts with limiting the direct effects of an attack, signaling attribution and ownership of an attack in a controllable manner, considering implications for other policy goals, on friends and allies, and even the Internet (Internet governance) itself, and finally the impact on international law and the evolving rules-based world order. Working with ambiguity is not only a feature of cyberattacks, but of cyber conflict writ large – therefore, not engaging in cyber operations (particularly in response) also impacts each of these elements as well. Nonaction is very much action in cyberconflict.

Ninth, SMPs now have a credible punishment capability but must be aware that the nature of escalation has changed – while armed conflict is an unlikely result, cross-domain escalation is indeed likely. Cyber operations have not been escalatory in the conventional military sense (leading to a use of force or armed attack) – an assumption that is the linchpin of the US persistent engagement doctrine's viability as an alternative to restraint strategies that rely on deterrence and explicit norms. While correct, this assumes a one-sided interpretation of escalation limited to armed conflict. Viewed within the broader domain of diplomacy, there can be multiple outlets for escalation. Ignoring this, offers a limited interpretation of escalation dynamics within current international relations, in which states increasingly use cyber, hybrid and non-military means of power below the threshold of conflict. A situation and relation can, and most often does, escalate in ways other than armed conflict. If we all agree that conflict today has taken hybrid forms of various military and non-military instruments of power, why do we cling on to a purely military logic of escalation? It is not inconceivable, and indeed quite appropriate in the present era, that escalation should find its way into areas other than the military. Furthermore, the same experts that confirm that such operations have rarely led to armed conflict so far, simultaneously warn that new US policies for authorizing preemptive offensive cyber strategies risk crossing thresholds and changing the rules of the game, and risk exacerbating a security dilemma through which each actor, convinced that they are acting defensively, reinforce a spiral of tit-for-tat escalation that could lead to armed force. SMPs should therefore be cognizant of the escalation dynamics and determine their objectives, whether they want to do full spectrum deterrence or build up a strategic limited deterrence capability, and act on it accordingly.

Tenth, SMP states are encouraged to tie deterrence efforts to more clearly demarcated thresholds, while allowing for strategic ambiguity to guard against hybrid tactics that deliberately seek to avoid such thresholds. This starts with articulating what malicious activity they seek to deter beyond the counterforce and countervalue critical assets that are considered out of bounds. They then need to signal their punishment capability in such a way that minimizes the risk of misinterpretation and inadvertent escalation. One way to do so is by exercises or by employing offensive tools that are clearly linked to redlines.

Annex I. The History of Deterrence Theory: a Primer

Deterrence is a form of behavior modification by one actor on another and can generally be defined as "the practice of discouraging or restraining someone (...) from taking unwanted actions. [...] It involves an effort to stop or prevent an action,"[217] usually achieved by "persuading an adversary that prospective costs would outweigh prospective gains"[218]. *Compellence* is a related term but differs slightly as it is described as "an effort to force an actor to do something."[219] Traditional deterrence is focused on punishment and can be summarized as "if you do this, I will do that", whereas compellence offers assurance in the form of "if you don't do this, I won't do that". While many variations of deterrence exist within the context of international affairs, it's useful to distinguish between *general* and *immediate* deterrence, as well as *direct* and *extended* deterrence.

General deterrence refers to normally stable situations where interstate relations are tense, but "the antagonism has long lost its edge,"[220] thus reducing the necessity of immediate deterrence.[221] In this case, the main objective is "to prevent short-term crises and militarized conflict from arising."[222] General deterrence is considered to be functioning when no one is planning an aggression against the status quo.[223] It largely coincides with what Huth describes as a situation where state representatives are prevented from making political or military threats during peacetime that could escalate to a crisis.[224] *Immediate deterrence*, comes into play when general deterrence has broken down,[225] describing circumstances where "antagonism is sharp,"[226], and one side is deliberating whether to "prepare for military action."[227] In this instance, the objective of immediate deterrence is "to prevent a specific, imminent attack."[228] For Huth, immediate deterrence refers to a situation where state

[217] Michael J. Mazarr, "Understanding deterrence" in *Deterrence in the 21st century—Insights from theory and practice*, eds. Frans Osinga & Tim Sweijs (Berlin: Springer, 2020), 15.

[218] Lawrence Freedman. "Introduction – The evolution of deterrence strategy and research," in *Deterrence in the 21st century—Insights from theory and practice*, eds. Frans Osinga & Tim Sweijs (Berlin: Springer, 2020), 5

[219] Mazarr, "Understanding deterrence," 15

[220] Freedman, "The evolution of deterrence strategy and research" 7

[221] Mazarr, "Understanding deterrence," 18

[222] Paul K. Huth, P. "Deterrence and international conflict: Empirical findings and theoretical debates." *Annual Review of Political Science, vol. 2*, no.1 (1999), 27.

[223] King Mallory, "New challenges in cross-domain deterrence," RAND Corporation, 2018, https://www.rand.org/pubs/perspectives/PE259.html, 3

[224] Huth, "Deterrence and international conflict," 27

[225] Mallory, "New challenges in cross-domain deterrence, 3

[226] Freedman, "The evolution of deterrence strategy and research," p. 7

[227] Mallory, "New challenges in cross-domain deterrence, 3

[228] Mazarr, "Understanding deterrence, 18

representatives have already escalated the situation into a crisis, and the goal then becomes to avoid the employment of their full weapon arsenal.[229] General deterrence can therefore be linked to the general armament of a nation, while immediate deterrence involves the deployment of those arms in a specific situation. Within the context of the Cold War, US general deterrence vis-à-vis the USSR included a strategy of issuing frequent public statements of retaliation threat in case of attack against the US territory or any territory of its allies.[230] Immediate deterrence came into play when rising tensions between the two superpowers could evolve into actual aggression,[231] for example the Cuban Missile Crisis of 1962.[232] For immediate deterrence, the stakes are higher as war might follow if the potential aggressor is not deterred.

Direct deterrence consists of a country's attempt to thwart attacks against its own territory[233] or threats to a country's vital interests by a potential military aggression froman enemy.[234] In *extended deterrence*, the defender is not trying to deter aggressive action against its own territory, but against the territory of third parties, such as allies. In situations of extended deterrence, the credibility of the threat is usually deemed lower because the deterrer is defending non-vital interests.[235] One well-known example of extended deterrence can be considered the creation of NATO and its Article 5.[236]

First wave - Early applications of deterrence theory (late 40s- early 50s)

Although limited, the first applications of deterrence theory to national security can actually be found before the Second World War, when strategic air raids were explored as a deterrent in the 1930s.[237] When faced with the rising power of Germany, British leaders "sought deterrence through offensive air power."[238] Early deterrence theorists focused on abstract logic instead of historical inquiry to analyze the behavior of state actors assumed to be rational.[239] They used deductive models and game theory to trace the cost-benefit calculations inherent to conflict decision-making. Deterrence pioneers like Snyder later built on the logic of these models.[240] According to Jervis' study *Deterrence Theory Revised*, these first analyses of nuclear deterrence were able to lay the groundwork for nascent scholars, even though the systemic approach that would follow in later waves was generally missing.[241] Additionally, early deterrence analyses had limited impact as they did not focus enough on the urgent international issues of the time as well as issues of national security[242].

[229] Huth, "Deterrence and international conflict," 27

[230] Mazarr, "Understanding deterrence," 18

[231] Ibid 18

[232] Arbatov, "Nuclear deterrence: A guarantee for or threat to strategic stability," in *Deterrence in the 21st century—Insights from theory and practice*, eds. Frans Osinga & Tim Sweijs (Berlin: Springer, 2020), 79

[233] Huth, "Deterrence and international conflict," 27. See also, Mazarr, "Understanding deterrence," 16

[234] Mallory, "New challenges and cross-domain deterrence," 3

[235] Mallory, "New challenges and cross-domain deterrence," 4

[236] Mazarr. "Understanding deterrence," 16

[237] Freedman, "The evolution of deterrence strategy and research," 3

[238] Jeffrey L. Hughes, "The origins of World War II in Europe: British deterrence failure and German expansionism," The *Journal of Interdisciplinary History, vol. 18*, no.4, 1988, 854.

[239] Scott D. Sagan, "Review: History analogy and deterrence theory," *The Journal of Interdisciplinary History, vol. 22*, no.1, (Summer, 1991): 79-88.

[240] Glenn Snyder, "Deterrence and power," *Journal of Conflict Resolution, vol. 4*, no. 2, (June, 1960): 163-178

[241] Jervis, "Deterrence theory revisited," 291

[242] Ibid,291

It was not until the emergence of nuclear capabilities at the end of WWII that deterrence theory came to the forefront in international relations debates. After witnessing the atomic bombings of Hiroshima and Nagasaki in 1945, early deterrence scholars began exploring the application of deterrence to nuclear weapons.[243] Scholars such as Bernard Brodie and Arnold Wolfers attempted to provide long-term reflections on nuclear weapons right after the introduction of the atom bomb in conflict.[244] In The Absolute Weapon (1946), Brodie provided key insights about the 'nuclear revolution', or the argument that nuclear weapons changed traditional ideas about warfare and governance. It changed the nature of warfare to the extent that the main purpose of a state's military forces shifted from winning wars to averting them.[245] This shift is tied to the destructiveness of nuclear weapons and the risk of mutual annihilation.

Though Brodie was not yet aware of the Soviet Union developing nuclear weapons at the time of his publication, he did correctly predict their ability to "produce them in quantity within a period of five to ten years."[246] In the immediate aftermath of WWII, large parts of the world, in particular Europe, were preoccupied with the rebuilding of their nations rather than developing nuclear weapons programs. They were discouraged to pursue a nuclear weapons program while peaceful nuclear energy programs were encouraged by the US.[247] Two notable European exceptions are France and the United Kingdom. Due to its early research and development of nuclear energy prior to WWII, France embarked on a nuclear weapons program during the 1950s and carried out its first successful nuclear test in the Sahara Desert of Algeria in 1960.[248] The UK's involvement in nuclear energy and weapon development dates back to 1940. Following WWII, the British leadership decided to pursue its own nuclear weapons program as a minimum nuclear deterrence, leading to its first nuclear test in 1952.[249] Outside of Europe, current nuclear weapon-owning states like India and Pakistan only began developing their nuclear power programs later on, during the second and third waves of deterrence theory.[250]

Second wave – Nuclear deterrence (1950s-1960s)

Faced with the power of nuclear weapons, scholars began to address deterrence in the context of developing a nuclear strategy.[251] The idea was "to understand how to best utilize the new power of nuclear weapons to deter Soviet aggression."[252] Second wavers, like

[243] Jeff Knopf, "The fourth wave in deterrence research," *Contemporary Security Policy,* (2010), 1

[244] Jervis, "Deterrence theory revisited," 291

[245] Bernard Brodie, *The absolute weapon* (New Haven: Yale Institute of International Studies, 1946), 76

[246] Brodie, "The absolute weapon," 63-69

[247] For example, the formation of the International Atomic Energy Agency (IAEA) in 1956: IAEA, "Statute of the International Atomic Energy Agency," 1956, https://www.iaea.org/about/statute. Also see: Giles Scott-Smith, "The necessary Marshall myth," *Clingendael Spectator*, (April, 2018), https://spectator.clingendael.org/en/publication/necessary-marshall-myth; Ko Colijn, "Einde verkrampte geheimhouding nederland over atoomwapens," Clingendael Spectator (May, 2020), https://spectator.clingendael.org/nl/publicatie/einde-verkrampte-geheimhouding-nederland-over-atoomwapens

[248] Nuclear Threat Initiative, "France nuclear overview," NTI, (May, 2016), https://www.nti.org/analysis/articles/france-nuclear/

[249] Nuclear Threat Initiative, "United Kingdom nuclear overview, (August, 2015), https://www.nti.org/analysis/articles/united-kingdom-nuclear/

[250] Raj Chengappa*, Weapons of peace: the secret sory of India's quest to be a nuclear* power (New Delhi: Harper Collins Publishers India, 2000)

[251] Knopf, "The fourth wave in deterrence research," 1

[252] Sagan "Review: History analogy and deterrence theory," 79

Thomas Schelling, Albert Wohlstetter and Glenn Snyder, built their research on deduction and game theory models.[253] Some examples are the chicken game, the prisoner's dilemma and the 'zero-sum game'. Generally, these scholars conceptualized deterrence success as making the other believe that one will not give in to their demands. Along these lines, deterrence does not expand on how an adversary's motives may be changed, but rather focuses on perceived threats.[254] Snyder included the important concepts of intent and credibility in his calculations on whether a potential aggressor will be deterred or pursue aggressive behavior.[255]

In his contribution to deterrence theory, Wohlstetter focused on what he calls "the balance of terror" between the two superpowers, where he explained the receding probability of total war is due to both sides fearing mutual nuclear annihilation.[256] This excludes limited war, or wars where the objectives, war zone and involved parties are restricted. He notes the difference between first-strike and second-strike capabilities. *First-strike* indicates the ability of a country to strike another country first, while *second-strike* refers to the counter-strike ability of a country.[257] He found the latter to be more stable than the former in terms of deterrence.[258] His findings were especially relevant to the nuclear strategies that were developed during those years with second-strike capabilities in mind.[259] According to Wohlstetter, "to deter an attack means being able to strike back in spite of it."[260]

In the context of the historical Sputnik launch in the late 1950s that increased fears of surprise attacks and ineffective air defenses, second wave scholars defined deterrence theory in terms of manipulation of threats to coerce the potential aggressor,[261] thus focusing more on deterrence by punishment than deterrence by denial.[262] However, from the perspective of non-nuclear small-to-medium sized countries, the focus was more placed on deterrence by denial and resilience. This was notably the case in Europe where, excluding the exceptions of France and Great Britain, states were not in possession of nuclear weapons to employ a minimum nuclear deterrent. Still, Europe was considered the most likely setting for a nuclear WWIII,[263] pushing several European countries, such as Switzerland, to seriously consider developing their own nuclear weapons program. Beyond building nuclear power plants for

[253] Sagan, "Review: History analogy and deterrence theory," 79. See also, Jervis, "Deterrence theory revisited," 291; Amir Lupovici, "The emerging fourth wave of deterrence theory – Toward a new research agenda," International Studies Quarterly, *vol. 54*, no. 3 (September, 2010): 705-732.

[254] Jervis, "Deterrence theory revisited," 292

[255] By going beyond the balancing of capabilities, traditionally a means of calculating relative power, calculating a 'balance of intentions' allows for tracing a more outlined strategy and a reduced dependence on military force. Decision-makers can calculate what portion of the adversary's available destructive power will be used in response to a range of their own moves. With regards to credibility, Snyder defines it as the "degree of probability that the power-wielder will actually carry out the threat if its terms are not complied with or will keep a promise if its conditions are met." Snyder, "Deterrence and Power," 164

[256] Albert Wohlstetter, "The delicate balance of terror, *Foreign Affairs, vol. 37*, no. 2 (January, 1958): 211-234

[257] Wohlstetter, "The delicate balance of terror," 211-234

[258] Jonathan Stevenson, "Thinking beyond the unthinkable: harnessing doom from the Cold War to the War on Terror," (New York: Viking, 2008). See also Lupovici, "The emerging fourth wave of deterrence theory," 705-732.

[259] Stevenson, "Thinking beyond the unthinkable"

[260] Wohlstetter, "The delicate balance of terror," 211-234

[261] Jervis, "Deterrence theory revisited," 292

[262] Lupovici, "The Emerging Fourth Wave of Deterrence Theory—Toward a New Research Agenda" https://www.deepdyve.com/lp/oxford-university-press/the-emerging-fourth-wave-of-deterrence-theory-toward-a-new-research-eRjBYhlklh?articleList=%2Fsearch%3Fquery%3Dthe%2Bemerging%2Bfourth%2Bwave%2Bof%2Bdeterrence%2Btheory%2Btoward%2Ba%2Bnew%2Bresearch%26docNotFound%3Dtrue)

[263] Yves Bouvier, "Nuclear fear in Europe: from weapons to power stations," *Encyclopédie d'histoire numérique de l'Europe*, (June 22, 2020) https://ehne.fr/en/encyclopedia/themes/material-civilization/risks-and-security/nuclear-fear-in-europe-weapons-power-stations

meeting national energy needs, Switzerland also studied the foundations for developing its own nuclear weapon in secret.[264] In the end, although it did make plans for underground tests and acquired uranium from the US, it cancelled the nuclear weapons program mostly for budgetary reasons.[265] The cost of doing research on and the testing, storing and safeguarding of nuclear weapons was deemed too high. An accident with a pilot reactor in 1969 increased the domestic opposition towards developing Swiss indigenous nuclear capabilities.[266] Countries like South Africa and Israel, also developed their own nuclear weapons in the 1960s with some degree of support or without opposition from the British, French and US governments.[267] Whereas the South African government was motivated to do so because of perceived threats by the Soviet Union, Israel faced existential threats from the surrounding Arab countries in the region. AlthoughIsrael still covertly maintains a nuclear arsenal, the South African government voluntarily dismantled its complete nuclear stockpile in the 1990s.[268]

While some nuclear patrons, most notably the US, initially encouraged some strategic partner countries to develop their own weapons program as a deterrent against the Soviet Union, they also became increasingly worried about the risks of such weapons being used and the reduced dependency such countries would have on the patron, thereby decreasing their leverage over the protégé.[269] A nuclear weapons sharing program with several European NATO allies, including the Netherlands, Belgium, Italy, Germany, Greece and Turkey, was conceived to prevent the further proliferation of nuclear weapons among the western allies.[270] By means of this program the US stored nuclear warheads in several airbases of its allies to provide a nuclear deterrent in relative proximity of the Soviet Union, many of which remain in place to date. As part of the conditions set to the sharing agreement, the use of these nuclear warheads is dependent on US authorization.[271] Therefore, while some Europeans seriously considered developing their own minimum nuclear deterrence capability, this was abandoned mostly because of budgetary reasons and external pressure. Instead, they relied on the extended nuclear umbrella of the US. It has to be noted that, fueled by doubts about the credibility of US extended deterrence,[272] some of these European coun-

[264] Jurg Stusi-Lauterberg, "Historical outline on the question of Swiss nuclear armament," (December, 1995), http://www.alexandria.admin.ch/bv001147186.pdf

[265] Gunnar Westberg, "Swiss nuclear bomb," (October, 2010), https://peaceandhealthblog.com/2010/10/09/swiss-nuclear-bomb/

[266] Marc Letau, "Swiss reactor meltdown", (March, 2019), https://www.swisscommunity.org/en/news-media/swiss-review/article/swiss-reactor-meltdown

[267] Nuclear threat initiative, "Fact sheet Israel nuclear overview," (October, 2021), https://www.nti.org/countries/israel/; Nuclear threat initiative, "Fact sheet South Africa nuclear overview," (October, 2019), https://www.nti.org/analysis/articles/south-africa-nuclear-disarmament/

[268] Nuclear threat initiative, "Fact sheet Israel nuclear overview," (October, 2021), https://www.nti.org/countries/israel/; Nuclear threat initiative, "Fact sheet South Africa nuclear overview," (October, 2019), https://www.nti.org/analysis/articles/south-africa-nuclear-disarmament/

[269] Jasen J. Castillo and Alexander B. Downes, "Loyalty, Hedging, or Exit: How Weaker Alliance Partners Respond to the Rise of New Threats," *Journal of Strategic Studies* 0, no. 0 (July 2020): 1–42, https://doi.org/10.1080/01402390.2020.1797690.

[270] Center for Arms Control and Non-Proliferation, "Fact sheet: US nuclear weapons in Europe," (August, 2021) https://armscontrolcenter.org/fact-sheet-u-s-nuclear-weapons-in-europe/

[271] Colijn, "Einde verkrampte geheimhouding Nederland over atoomwapens?"

[272] The extent of these doubts can be exemplified by French President Charles de Gaulle disbelieving the US willingness to intervene in case of aggression in Europe. He is known for having questioned US President Kennedy on whether he would actually be willing to trade New York for Paris. This debate clearly shows how uncertain US allies were of US promises and their willingness to actually enter in a nuclear war with the USSR in case of aggression against a European country. Eventually, this lack of credibility pushed France to acquire its own nuclear weapons. Mazarr, "Understanding deterrence," 16; Orion Noda, "Risking New York for Paris? The illusion of the US nuclear umbrella," (May, 2020), https://www.strifeblog.org/2020/05/01/7577/

tries turned their attention towards resilience and denial measures through the widespread construction of underground bunkers.[273]

Second wave deterrence theorists came under heavy criticism, primarily for the lack of empirical data from the employed methodology. According to subsequent scholars, they relied too much on deductive reasoning and hypothesized scenarios without taking into account historical case studies.[274] Moreover, their analysis did not clarify whether deterrence theory actually influenced the behavior of decision-makers and arguably placed too much emphasis on commitment, rather than interest, when analyzing the losses of a state being deterred.[275] Additional criticism includes a lack of consideration of non-military elements of deterrence, such as deterrence through the promise of rewards and through compromises, and not addressing significant limitations to the theory. This refers, for example, to the adoption of an ethnocentric approach whereby it was assumed that countries might differ in their goals, but not their view of the world.[276] Ultimately, this does not allow for a proper understanding of the enemy's values and motivations.[277] Related to this point, scholars like Jervis believed there to be too much focus on the rationality of decision-makers.[278] He identified four "barriers to accurate perception" which limited a decision-maker's ability to respond to unexpected situations: overconfidence, not seeing value trade-offs, the assimilation of new information to pre-existing beliefs and defensive avoidance.[279] Additionally, when decisions are made in the context of a crisis involving nuclear weapons, factors like time-pressure, risk tolerance, incomplete or even conflicting intelligence and psychological stress play a large role.[280]

Second-strike nuclear capabilities also led to the concept of 'mutually assured destruction' (MAD) – when an attacker's first use of a nuclear weapon, would be reciprocated by the defender. It is said to have influenced the de-escalation of conflicts where both parties were in the possession of a nuclear arsenal, like the USSR vs. the US and India vs. Pakistan. Within the context of the Cold War, the two superpowers realized that nuclear escalation would have led to mutual destruction. Thus a stable, albeit fragile, situation prevailed. This consideration was shaped especially by the 1962 Cuban Missile Crisis,[281] where the US and USSR came so close to actual war that it provided the necessary push to start seeking for common grounds, such as the 1963 Partial Test Ban Treaty and the 1968 Treaty on the Non-Proliferation of Nuclear Weapons.[282]

As a result, Schelling moved away from the zero-sum considerations that had dominated the debate until then. Instead, he argued that the two superpowers could effectively identify

[273] For example, starting from the 1940s until the early 2000s, Sweden constructed 65,000 fallout shelters to protect seven million people in case of war. Similarly, from the 1960s, Switzerland began to introduce legal requirements whereby each citizen must have a protected place in a shelter located near its house. Catherine Edwards, "Why Sweden is home to 65,000 fallout shelters," (November, 2017), https://www.thelocal.se/20171101/why-sweden-is-home-to-65000-fallout-shelters/. Also see: Daniele Mariani, "A chacun son bunker," (October, 2009), https://www.swissinfo.ch/fre/a-chacun-son-bunker/7485678; BBC, "The bunkers built to survive an apocalypse," BBC, (August, 2017) https://www.bbc.com/future/article/20170825-the-bunkers-built-to-survive-an-apocalypse;

[274] Sagan, "History, analogy and deterrence theory," 79

[275] Jervis, "Deterrence theory revisited," 314

[276] According to Jervis, theorists tended to assume that different countries had the same vision of the world, thus not taking into account the possibility of people constructing differing geopolitical analyses. See Jervis 1979, p. 296

[277] Jervis, "Deterrence theory revisited," 296-297

[278] Ibid, 294-301

[279] Ibid.

[280] Stevenson, "Thinking beyond the unthinkable"

[281] Arbatov, "Nuclear deterrence," 81

[282] Ibid.

shared interests,[283] and that "the mutual release of some information through bargaining could lead to more agreeable outcomes for both sides."[284] The shift away from abstract calculations of deterrence towards a more practical need for cooperation is particularly visible when looking at escalation. Until the 1960s, the concept of escalation indicated that any critical conflict between the two nuclear powers would automatically lead to total nuclear war, thus making it inevitable.[285] A reconsideration of escalation allowed for it to be understood less as an inescapable event, and rather as a potential outcome leading to total war.[286] Kahn's Nuclear Escalation Ladder (1965) is a telling example.[287] It consists of 44 steps of conflict escalation illustrating how a crisis could surpass various thresholds and reach the highest rung by developing into nuclear war, or "insensate war."[288] His idea of escalation dominance assumes that deterrence would be guaranteed by a state's military superiority at each level of escalation.[289] In this way, according to Schelling, deterrence would be stabilized as the weaker party would avoid escalation to save face.[290] A review of the inevitability of nuclear war was in line with the superpowers' strategies at the time, which focused more on flexible responses.[291] Moreover, Kahn's ladder, specifically, became to be considered a practical tool to inform nuclear strategy and used "to define where escalatory action could be arrested or controlled" to avoid nuclear war.[292]

Third wave – Conventional deterrence (1970s-1980s)

While the two superpowers were finding ways to de-escalate, deterrence theorists focused their analysis on how to reduce the chances of nuclear war by stabilizing the global political environment.[293] Informed by the détente of 1969-1979, deterrence theory underwent several adjustments, pushed forward by third wave scholars.[294] Firstly, the concept of deterrence by denial started to gain more popularity. Previously, Snyder already asserted that the latter would be the better option for decision-makers, promising a less escalatory means of defense.[295] Snyder also highlighted non-military means of deterrence through

[283] Stevenson, "Thinking beyond the unthinkable"

[284] Ibid.

[285] Freedman, "The evolution of deterrence strategy and research",3

[286] Ibid.

[287] Herman Kahn, "*On Escalation: Metaphors and Scenarios*," NY Praeger, 1965 (republished by Transaction Publishers, New Brunswick, N.J., 2010, with a new foreword by Thomas C. Schelling).

[288] Rodney Jones, "Nuclear Escalation Ladders in South Asia." *Policy Architects International*, (2011), 31-322

[289] Stevenson, "Thinking beyond the unthinkable"

[290] Stevenson, "Thinking beyond the unthinkable"

[291] Jones, "Nuclear Escalation", 5

[292] Ibid.

[293] Stevenson, "Thinking beyond the unthinkable"

[294] The easing or relaxation of tensions through mutual understanding, compromise and improved diplomacy between the US and Soviet Union. Some scholars argue that détente began (and continued longest) in Europe, because of a relaxation of East-West tensions in Germany under Chancellor Willy Brandt and the conclusion of the Conference on Security and Cooperation in Europe (CSCE) in 1975. Along with increased competition and the failed implementation of the SALT II arms control agreement, the Soviet intervention in Afghanistan of 1979 brought an end to the détente period.

[295] Snyder, "Deterrence and defense", 15

trade restrictions and the promise of rewards.[296] It reopened the focus on conventional deterrence.[297] It is more broadly defined as "all deterrence that doesn't involve threats to use nuclear (or other unconventional) weapons,"[298] or more narrowly as "deterrent threats to resist or to inflict costs against an aggressor using conventional military force during the resulting conflict."[299] According to Freedman, the increasing attention on deterrence by denial allowed NATO to play a more active part in its member states' deterrence strategies, while relying on the American nuclear umbrella to promote deterrence by punishment.[300]

Whereas the second wave predominantly focused on theorizing a nuclear strategy based on deterrence theory, third wavers moved away from deterrence through the direct threat of a nuclear strike to nuclear weapons being considered a last-resort option.[301] Here, the threat and risk of nuclear escalation, while still possible, became less probable.[302] By focusing on real-life examples of conventional deterrence, they were able to re-analyze past incidents using statistical data and provide empirical findings on what elements work in successful deterrence theory.[303] It demonstrated that deterrence theory needed to incorporate aspects of taken risk, rewarding, probability, misperceptions and domestic political circumstances.[304]

While analyzing the findings of the second wave, Jervis identified two main shortcomings in contemporary deterrence theory to which he proposed various modifications. Firstly, he believed that up to that time, deterrence had been approached with an apolitical stance, and little consideration of the context within which measures, threats and events occur.[305] This same critique was pushed forward by Knopf a few decades later, when he recognized that "deterrence must be tailored to each individual case based on a detailed understanding of the other side."[306] Secondly, Jervis supported the notion that a state's cost of retreating, or the values a party is sacrificing, are more dependent on a party's intrinsic interests, rather than commitment.[307] Accordingly, interests vastly influence the deterrence process, as the party with the higher intrinsic interest will have higher costs of retreating, a higher chance of prevailing, and higher chances to succeed in the bargaining process.[308]

This critique was especially important to propose new elements of what constitutes deterrence success. In his empirical study, Huth asserted that deterrence may be more likely to succeed when a defender can deny a potential attacker a quick and decisive victory and there is "a policy of reciprocity in diplomacy and military actions by the defender" (tit-for-tat).[309]

[296] Snyder, "Deterrence and power", 163

[297] According to US doctrine, conventional weapons can be understood as all weapons excluding those of mass-destruction, including nuclear, biological and chemical. Department of Defense Dictionary of Military and Associated Terms, "Joint Pub 1-02 or JP 1-02," (April, 2001 As Amended Through April, 2010), 106 https://web.archive.org/web/20100814230117/http:/www.dtic.mil/doctrine/new_pubs/jp1_02.pdf

[298] Mueller Carl, "The Continuing Relevance of Conventional Deterrence", 50

[299] Ibid.

[300] Freedman, "The evolution of deterrence strategy and research", 5

[301] Freedman, "The evolution of deterrence strategy and research", 6

[302] Ibid.6

[303] Jervis, "Deterrence theory revisited", 301. See also, Freedman, "The evolution of deterrence strategy and research", 5. See also, Knopf, "The fourth wave in deterrence research", 1

[304] Ibid, 303-314

[305] Ibid, 323

[306] Knopf, "The fourth wave in deterrence research", 9

[307] Jervis, 1979, p. 314-315; He defines intrinsic interests as "the inherent value that the actor places on the object or issue at stake", which differs from commitment which is, instead, "manipulated by the state to increase its costs of retreating and thereby improve its bargaining position"

[308] Jervis, "Deterrence theory revisited"

[309] Paul Huth, "Extended Deterrence and the Outbreak of War," *American Political Science Review, vol.* 82, no. 2 (June, 1988) pp. 423

Third wave theorists emphasized the importance of taking context and the role of perception into consideration: "deterrence turns out to be much more than merely threatening a potential adversary: it demands the nuanced shaping of perceptions so that an adversary sees the alternatives to aggression as more attractive than war."[310] In his analysis, Jervis views deterrence success as proportional to the ability of decision-makers to understand the other side's perceptions.[311] He argues that misperception of the opponent's values and fears might lead to a misconstructed deterrence strategy and thus to its consequent failure.[312] Furthermore, misperception can also lead to self-deterrence when states perceive inexistant threats, and are therefore "deterred by the figments of their own imagination."[313]

Fourth wave – Non-state capabilities and actors (1990s-)

With the fall of the Berlin Wall came the fourth wave of deterrence theory. The geopolitical reality of only two nuclear superpowers was transformed into a variety of actors being considered threats to national security. The sources of threat became much more dispersed than before: non-state terrorists following the 9/11 attacks,[314] states from the second nuclear age (India, Pakistan and Israel) as well as rogue countries, such as Iraq, Iran and North Korea. Due to the changing nature of conflict, which now predominantly takes place below the threshold of war, relies on non-conventional methods, most notably cyber and information operations. This shift was driven by the imbalance of conventional capabilities between state and non-state actors. The military superiority of major states such as the US could not be matched by new non-state and state actors, so the latter developed asymmetrical tactics to avoid direct confrontation.[315] The application of deterrence strategy to these new hostilities resulted in attempts to reduce the number of attacks, instead of avoiding any type of conflict,[316] and to several other developments, most notably cross-domain deterrence. With the development of new capabilities, deterrence has been expanded to encompass potential conflict taking place in and across other, less-traditional domains (such as space, hybrid warfare, non-state actors, and cyberspace).[317] Successful *cross-domain deterrence* is present "when an opponent has no incentive to initiate or escalate conflict at any given intervention or escalation threshold in any given domain of warfare – both vertically and horizontally into one or more additional domains of warfare."[318]

The breakthrough of new non-conventional capabilities led to the birth of cross-domain deterrence. Different from the Cold War period, conflict began to be envisioned in domains that lacked a physical and geographical, dimension.[319] Kahn's escalation ladder, which had been applicable during the Cold War context, was deemed less relevant to the security and

[310] Mazarr, "Understanding deterrence", 15

[311] Reynold Jervis, "Deterrence and Perception," *International Security, vol. 7*, no. 3 (Harvard University Press, Winter, 1983), 3

[312] Ibid.

[313] Ibid, 14

[314] 9/11 demonstrated that both state and non-state actors developed resources to significantly strike a nation. Freedman,"The evolution of deterrence strategy and research", 8. See also, Knopf, "The fourth wave in deterrence research", 2

[315] Mallory, "New challenges in cross-domain deterrence", 1

[316] Knopf, "The fourth wave in deterrence research", 4

[317] Mallory, "New challenges in cross-domain deterrence", 6

[318] Ibid, 1

[319] Ibid, 6

political environment at the time.[320] This steered decision-makers towards the analysis on how hostile actors in subareas of warfare, such as hybrid conflicts, space, cyber and terrorism.[321] Scholars created a cross-domain escalation path, showing the potential route of a conflict escalating, both vertically – by crossing various thresholds of increased violence – and horizontally – by changing domains, such as from diplomatic to economic to the ultimate escalation threshold of nuclear warfare.[322]

Whereas Kahn placed nuclear annihilation as the highest step on his escalation ladder, alternative models are said to better explain the increasing interlinkage of the different domains in general (primarily diplomatic, military, economicm and informational) as well as specific military domains (conventional, nuclear, cyber, space) that characterize the 21st century. Rather than a linear ladder of escalation, the alternative escalation 'vortex' allows for a visualization of escalations in the different domains and their effect on each other. When escalatory ladders overlap (different means produce ends in the same domain), the outcome is that states produce nonlinear responses to actions.[323] Actions that are typically associated with one domain, for example cyber, can have rigorous effects in the conventional, nuclear and spatial domains. The escalation vortex, therefore, also provides insight as to why rule-based approaches to managing escalation pertaining to just one domain are not as successful.[324] Instead, the capabilities that actors have to respond to different levels of provocation in each domain can be represented, which can illustrate vulnerabilities or potential opportunities when developing national security strategies to manage escalation.[325] The escalation vortex is especially interesting for smaller states who have limited (offensive) capabilities to still place themselves in an advantageous position if they efficiently fill the gaps in their deterrence strategy across different domains.

Research on asymmetric violence and non-state actors also provided ways to understand how traditional aspects of deterrence could be applied to these new domains and capabilities. In his analysis, Mallory analyzes how traditional responses to potential aggressions, namely deterrence by denial and deterrence by punishment, can be applied to space, hybrid warfare, terrorism and cyber.[326] He found that, generally, deterrence by denial is a better strategy compared to deterrence by punishment, as the former only requires control while the latter necessitates constant coercion.[327] In 'gray-area' warfare, Freedman notes that denial-based deterrence is more appropriate as well.[328] Unless specific actions can be tied to specific actors, punishment-based deterrence will not be as effective.

The expansion of the application of deterrence, its methodologies and its targets allowed scholars to explore whether or not unconventional actors, such as terrorists, can effectively be deterred.

In his work, Lipovici presents both sides of this debate. On the one hand, scholars have found that terrorism is deterrable as long as there is a degree of overlap between the deterrer and

[320] R J, Vince, "Cross-Domain Deterrence Seminar Summary Notes", Lawrence Livermore National Laboratory (May, 2015), https://cgsr.llnl.gov/content/assets/docs/SummaryNotes.pdf

[321] Mallory, "New challenges in cross-domain deterrence", 2

[322] Ibid.

[323] Vince, "Cross-Domain Deterrence Seminar Summary Notes"

[324] Ibid.

[325] Ibid.

[326] Mallory, "New challenges in cross-domain deterrence", 8-19

[327] Ibid, 21

[328] Freedman, "The evolution of deterrence strategy and research", 8-9

the terrorist's preferences.[329] On the other hand, other scholars assume that, because terrorists may want to be targeted by nation states to legitimize their ideology, terrorism cannot be deterred.[330]

One suggested means to deter terrorism is through 'indirect deterrence.' Indirect deterrence is targeted against "third parties whose actions could affect the likelihood that a potential attacker can or will carry out an attack,"[331] thereby seeking to cut off their support and resources. Although this theory leans more towards the notion of compellence, states supporting or facilitating terrorist activities could be held accountable by indirect deterrence.[332] A similar application of deterrence tries to utilize forensics as support. By observing WMD-seeking terrorists, scholars suggested that state attribution could be identified through nuclear forensics, through the analysis of nuclear debris and tracing it back to its producing country.[333] Additional innovative applications of deterrence include deterrence by counternarrative and cumulative deterrence. Deterrence by counternarrative aims at delegitimizing terrorist groups vis-à-vis their communities.[334] While cumulative deterrence, stemming from the Israeli context, argues that terrorism could be deterred by a combination of factors, namely stopping individual terrorists, retaliating against successful attacks and long-term patience.[335] Eitan Shamir adds that by including the special characteristics that violent non-state actors harbor, fourth wavers have stretched the original concept of deterrence beyond its boundaries.[336]

The fourth wave of deterrence theory focused on new capabilities and new actors to define whether deterrence could still be applicable in a multi-polar world. Some fourth wave scholars signal a shift from deterrence as a distinct strategy to deterrence as part of a broader influence strategy to "increase the costs of attacks as well as rewards to increase the benefits of restraint simultaneously."[337] Due to the difficulties associated with accurate attribution with the use of non-conventional capabilities and with adequate retaliation vis-à-vis non-state targets, the focus of deterrence shifted from deterrence by punishment to deterrence by denial.[338] Nevertheless, scholars at the time came to the general conclusion that innovative applications of deterrence could also be utilized to deter hybrid hostilities.

[329] Amir Lupovici, "The Emerging Fourth Wave of Deterrence Theory — Toward a New Research Agenda,"), 718

[330] Ibid.

[331] Knopf, "The fourth wave in deterrence research", 11

[332] Ibid.

[333] Knopf, "The fourth wave in deterrence research", 20

[334] Ibid, 18

[335] Ibid, 14

[336] Eitan Shamir, "Deterring Violent Non-state Actors," in F. Osinga and T. Sweijs (eds.), NL ARMS Netherlands Annual Review of Military Studies 2020, NL ARMS, (2020), 283

[337] Peter V. Jakobsen, "Deterrence in Peace Operations: Look Beyond the Battlefield and Expand the Number of Targets and Influence Mechanisms," in F. Osinga and T. Sweijs (eds.), *NL ARMS Netherlands Annual Review of Military Studies 2020,* (2020), 332-335

[338] Freedman, "The evolution of deterrence strategy and research", 9

Annex II. Intelligence agencies

As noted in Chapter 3, the game of offense remains dominated by the intelligence community and hybrid – often non-state – actors that often operate with a direct or indirect link to government. This Annex provides a more detailed description of the role that intelligence agencies play in cyber operations by the US, UK, CN, RF.

Within the US, the NSA Tailored Access Operations (TAO) unit has been and remains one of the most notable players in the field, although USCYBERCOM has arisen as the operational command in charge of offensive cyber operations. TAO was aggressively expanded to develop a global architecture and tools that can augment its traditional passive intelligence collection into covert action and sabotage.[339] As noted by one of the authors, "with TAO, SIGINT was suddenly being redefined 'in terms of breaking and entering', much more the CIA remit, and the term 'SIGINT at rest' (as opposed to "SIGINT" in motion) was coined to allow this type of activity, which included both hacking a computer and the physical plating of a listening device".[340] This extended to preparation of the battlefield (OPE), as shown by operations like Nitro Zeus, the follow-up operation to the Olympic Games Operation, which includes Stuxnet, in case the Iran nuclear dispute led to a conflict. The New York Times reported that "while Cyber Command would have executed Nitro Zeus, the National Security Agency's Tailored Access Operations unit was responsible for penetrating adversary networks, which would have required piercing and maintaining a presence in a vast number of Iranian networks, including the country's air defenses and its transportation and command control centers."[341]

In fact, several reports and leaks have indicated that the Flame malware, first reported by Kaspersky in 2012, was actually of US design and potentially part of the Olympic Games Operation. Worryingly, it featured hijacked Microsoft certificates that lies at the very core of trust-relations on the Internet.[342] It shows a willingness to take high political risks in the

[339] Traditionally, the US strategies and budgets appear to favor offensive over defensive. Programs that appeared as purely defensive to the outside, actually heavily leaned towards offensive measures, especially the capabilities of the NSA. See for example the 2008 Comprehensive National Cybersecurity Initiative (CNCI) that initially appeared to be purely defensive and made reference that would make outsiders consider it to be offensive. From the Snowden leaks, however, it became clear that it served as the direct justification for an annual $650 million USD expansion of NSA operations direct at securing presence on endpoints. TAO's endpoint activities largely rest on obtaining remote access via covert implants and infrastructure throughout the world. To illustrate, in 2004, NSA was managing about 100–150 implants worldwide to 21,252 by 2008 and was projected to control 85,000 by the end of 2013. Klimburg, *The Darkening Web,* 153; Ryan Gallagher and Glenn Greenwald, "How the NSA Plans to Infect 'Millions' of Computers with Malware", The Intercept, (12 March, 2014) https://theintercept.com/2014/03/12/nsa-plans-infect-millions-computers-malware/; and Barton Gellman and Ellen Nakashima, "U.S Spy Agencies Mounted 231 Offensive Cyber-Operations in 2011, Documents Show," Washington Post, (August 30, 2013), https://www.washingtonpost.com/world/national-security/us-spy-agencies-mounted-231-offensive-cyber-operations-in-2011-documents-show/2013/08/30/d090a6ae-119e-11e3-b4cb-fd7ce041d814_story.html

[340] Klimburg, *The Darkening Web,* 172. General Hayden – former chief of the NSA and USCYBERCROM – also describes that "we pretty much had all we needed to thrive, at least in terms of law and policy" and marveled how NSA remarkably and silently "went from a world of letting radio waves serendipitously hit our antennas to what became a digital form of breaking and entering". Michael V. Hayden, *Playing to the Edge: American Intelligence in the Age of Terror* (New York: Penguin Books, 2016)

[341] Sanger and Mazzetti, "U.S. Had Cyberattack Plan if Iran Nuclear Dispute Led to Conflict,"

[342] Klimburg, *The Darkening Web,* 184

pursuit of intelligence gains, indicative of the American path dependency in cyber operations in which the bottom-up approach places the tactical before the political. In other words, there is a natural trend favoring the technical operators in a highly complex and esoteric field as cyber and a "bottom-up culture of putting technical feasibility before political desirability, which is hardwired into the NSA and US Cyber at large."[343] It also shows that there was a complete absence of public discussion and relatively few constraints and congressional scrutiny on NSA's foreign intelligence activities compared to domestic ones. The normative and legal context may be changing, in part because of the blurring between the traditional, foreign and domestic distinctions governing intelligence activities, and as result of increased public awareness and foreign scrutiny of the 'by any legal means' necessary mantra of the US. Given the clandestine nature of the internal workings of the intelligence apparatus and its operations, it is difficult to determine how these legal and normative questions are navigated internally and to what extent they affect the bottom-up problem of placing the tactical before the political so illustrative of American cyber operations. That normative and legal context may be changing, not just because Internet routing increasingly blurs the traditional foreign and domestic distinction governing intelligence activities, but also as result of increased public awareness and foreign scrutiny of the 'by any legal means' necessary mantra of the US. Given the clandestine nature of the internal workings of the intelligence apparatus and its operations, it is difficult to determine how these questions are navigated internally and to what extent they have changed the American path dependency.

Other countries, like the UK decided to embed their wide range of cyber operations, from defensive, to intelligence and offensive, in their SIGINT agencies, namely the Government Communications Headquarters (GCHQ). Together with the Ministry of Defense, it jointly runs the National Offensive Cyber Program, which is reported to have a budget £250 million and a staff of 2,000 in 2018.[344] The Snowden leaks revealed four classified documents that showed GCHQ has a covert cyber warfare unit, labelled the Joint Threat Research Intelligence Group (JTRIG). It was previously described to engage in online covert action, mainly through information operations and technical disruptions to "deny, disrupt, degrade, deceive". Furthermore, its Human Science Operations Cell also engages in online covert action that focuses on online human intelligence, strategic influence and disruption and computer network attacks with the aim of understanding and manipulating the wider online discourse.

Within China, where the military is the main actor, a second category of PLA-authorized forces in civilian organizations can be identified. These include the foreign-intelligence Ministry of State Security (MSS) and the domestic Ministry of Public Security (parts of which took their orders from the Chinese Cyber Administration). Economic espionage activities underwent changes in command and control – away from PLA and towards MSS – and in apparent levels of sophistication, operating more clandestinely with more sophisticated tactics, techniques and procedures.[345]

But in no other country is the role of the intelligence community in carrying out cyber operations so prominent as in Russia. The respective parties include the Main Directorate of the General Staff of the Armed Forces (GRU), the Foreign Intelligence Service (SVR), the Federal Security Service (FSB) and the Federal Protective Service (FSO). First, the GRU oversees

[343] Ibid,149-150

[344] The Telegraph, "Britain steps up cyber offensive with new £250m unit to take on Russia and terrorists;" Telegraph.co.uk, (September 21, 2018), https://www.telegraph.co.uk/news/2018/09/21/britain-steps-cyber-offensive-new-250m-unit-take-russia-terrorists/

[345] This change in sophistication becomes apparent in the reporting of outside (mainly US) cybersecurity companies and US government indictments of Chinese hackers.

both strategic and tactical level intelligence collection and is most well-known for carrying out "Russia's most brazen and damaging cyberattacks", including the 2015 attack against Ukraine's electrical infrastructure, the 2016 US presidential election and the 2017 French presidential election, the 2017 NotPetya ransomware attack, and the 2018 hacking attempt against the Organisation for the Prohibition of Chemical Weapons (OPCW) amongst many others.[346] It is difficult to compare to agencies from other countries as it not only houses usual forms of intelligence and cyber operations (unit 26165 and 74455) but also disinformation units (Unit 54777) and the large special-operations forces used for peacetime covert action as well as wartime sabotage missions (the infamous *Spetsnaz* troops).[347] It is considered to have the best technological capabilities among all Russian special services. Second, the SVR is the primary civilian foreign intelligence agency that mostly targets government networks, think tanks, and information technology companies, and was found to be behind the SolarWinds operation.[348] APT29 – also known as CozyBear or the Dukes – has been linked to the SVR by the Dutch intelligence services and cybersecurity company CrowdStrike. Third, the FSB is Russia's leading domestic security agency, responsible for counterintelligence to protect the nation from foreign cyber operations.[349] It is the principal successor to the KGB and when the former SIGINT agency (FAPSI) was dismembered, the FSB was the main recipient of its expertise and also its crown jewel: SORM or the System of Operational-Investigative Activities that allows it to monitor all telecom traffic within Russia. Not much is known about the various FSB cyber units, except for cursory evidence pointing to a number of cyber espionage groups associated with it, including APT29, which is also linked to the SVR. Fourth, the FSO, in particular the Spetssvjaz sub-unit, is considered to exclusively focus on defending the physical and electronic security of government and military institutions, communications and personnel through SIGINT and electronic capabilities. The FSO is considered to be "an overseer of the various security services, helping to monitor infighting and the accuracy of intelligence report".[350] While some of the Russian intelligence agencies are more focused on building their internal capabilities than others, the most defining organizational characteristic is the hybrid nature in which they operate closely with outside contractors, cybercriminals and hacktivists.

[346] Congressional Research Service, "Russian Cyber Units," CRS Reports, (January 4, 2021) https://crsreports.congress.gov/product/pdf/IF/IF11718

[347] From its mission-specific directorates, the Sixth Directorate, in charge of electronic/signals intelligence, hosts the notorious Unit 26165 and Unit 74455. Unit 26165 was established as the 85th Main Special Service Center responsible for military intelligence cryptography during the Cold War. Unit 74455, on the other hand, appears to be a much more recently-established unit to help expand GRU cyber capabilities. Also known as the Main Center for Special Technologies or for outsiders as *Sandworm*, the Unit was indicted by the DoJ for a number of cyber operations. Finally, there is Unit 54777, known as the 72nd Special Service Center, that is responsible for psychological operations, more recently linked to online disinformation campaigns targeting the Covid-19 pandemic. They provide support to the other cyber units and also operate on the tactical level by harmonizing electronic warfare and information warfare operations.

[348] National Cyber Security Centre, "UK and US call out Russia for SolarWinds compromise," https://www.ncsc.gov.uk/news/uk-and-us-call-out-russia-for-solarwinds-compromise

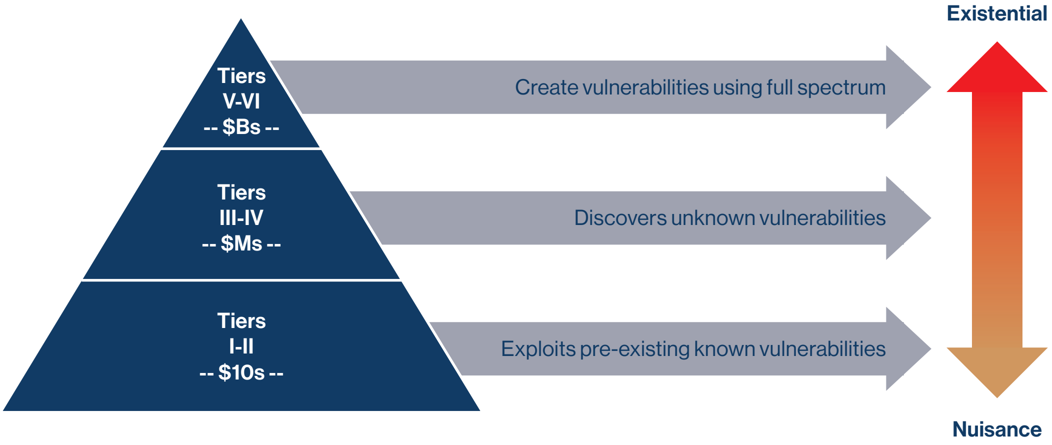
[349] The 16th and 18th Centre of the FSB are its main signals and cyber intelligence units. Together with Department K of the Ministry of Interior, it also monitors domestic criminal hackers, and more recently its exclusively domestic area of operations (the 18th Centre in particular) has reportedly expanded to foreign spheres of influence, most notably neighboring post-Soviet states, causing friction with the GRU and SVR.

[350] Congressional Research Service, "Russian Military Intelligence: Background and Issues for Congress"

Annex III. The tiers of offensive cyber capabilities

There are numerous indices that categorize nations across different tiers of power. The tiered categorization used in this report is largely based on the framework designed by the Defense Science Board of the US Department of Defense.[351] Other categorizations have emerged since, such as the IISS' "Cyber Capabilities and National Power: A Net Assessment" and the Harvard Belfer Center's "National Cyber Power Index". Both assess the offensive capabilities of a number of states but also take a broader approach to measuring cyber power – e.g. by including the regulatory regime or development of norms.[352] The taxonomy adopted by the Defense Science Board, however, focuses on parsing the offensive capability of states along three categories: "those practitioners who rely on others to develop the malicious code, those who can develop their own tools to exploit publicly known vulnerabilities as well as discovering new vulnerabilities, and those who have significant resources and can dedicate them to creating vulnerabilities in systems".[353] Figure 3 visualizes this taxonomy, wherein the dollar figure indicates the nominal investment required to participate at the given tier. The pyramid shape suggests the decreasing number of practitioners as one ascends from the lower tiers, mostly described as a nuisance, to the highest tiers, which can pose an existential threat.

Figure 3. Cyber Threat Taxonomy, Defense Science Board (2013)


Existential
Tiers V-VI -- $Bs --
Create vulnerabilities using full spectrum
Tiers III-IV -- $Ms --
Discovers unknown vulnerabilities
Tiers I-II -- $10s --
Exploits pre-existing known vulnerabilities
Nuisance

[351] US DoD Defense Science Board, "Task force report: Resilient military systems and the advanced cyber threat," US Department of Defense (January, 2013)

[352] IISS, "Cyber capabilities and national power: A net assessment." IISS Research Paper, (June 28, 2021), Cyber Capabilities and National Power: A Net Assessment (iiss.org); Julia Voo et al., "National Cyber Power Index 2020", Harvard Kennedy School Belfer Center for Science and International Affairs (September, 2020), https://www.belfercenter.org/sites/default/files/2020-09/NCPI_2020.pdf.

[353] US DoD Defense Science Board, "Task force report: Resilient military systems and the advanced cyber threat," 21

The Defense Science Board further defines each Tier as follows:[354]

Tier	Description
I	Practitioners who rely on others to develop the malicious code, delivery mechanisms, and execution strategy (use known exploits).
II	Practitioners with a greater depth of experience, with the ability to develop their own tools (from publicly known vulnerabilities).
III	Practitioners who focus on the discovery and use of unknown malicious code, are adept at installing user and kernel mode root kits10, frequently use data mining tools, target corporate executives and key users (government and industry) for the purpose of stealing personal and corporate data with the expressed purpose of selling the information to other criminal elements.
IV	Criminal or state actors who are organized, highly technical, proficient, well-funded professionals working in teams to discover new vulnerabilities and develop exploits.
V	State actors who create vulnerabilities through an active program to "influence" commercial products and services during design, development or manufacturing, or with the ability to impact products while in the supply chain to enable exploitation of networks and systems of interest.
VI	States with the ability to successfully execute full spectrum (cyber capabilities in combination with all of their military and intelligence capabilities) operations to achieve a specific outcome in political, military, economic, etc. domains and apply at scale.

The Defense Science Board also makes three notes regarding the higher-tier actors. First, they still use the techniques at the lowest level to avoid exposing their more sophisticated techniques. Second, states, especially those that have fluid relations with non-state actors, might employ non-state actors as proxies, allowing middle-tier organizations to gain access to higher-tier capabilities. In turn, this blurs the distinctions made in this taxonomy. Third, the primary distinctive feature between Tier V and VI is the scale at which an organization can execute cyber operations that are highly complex and require a lengthy preparation. The discriminator of a Tier VI actor is therefore "funding, people and equipment to conduct many such operations concurrently."[355]

Transposing this taxonomy of tiers to state capabilities, including for attack (CNA) espionage (CNE) and defensive (CND), five classes are identified in the table below.[356] SMPsreferred to in this report show a medium or high degree of digitalization, medium to very high cyber defense with variations in readiness, and limited to specialized offensive cyber capabilities that span from Class 3 to 5 (or Tier II to V).

[354] Ibid, 22.

[355] Ibid, 23.

[356] Derived from: Klimburg (2013), edited from "Swiss MoD Hearing on Cyber Threats".

Cyber Capability Levels

Power	No.	Attributes	Prob. CNA	Prob. CNE	CND
Class 1 (Tier VI) "hyper cyber power"	USA	Dominant diplomatic position, highly developed military intel cap integrated with private sector, deployed as part of DIME, but relatively poor defense	Highly developed military (battlefield support), SIGINT gathering, OCEO inc. IO = IW?	Most extensive, heavy SIGINT / 3rd party SCS	Medium to high for gov. mil systems, medium-to-poor for "national" systems – focus on "deterrence"
Class 2 (VI) "major cyber power"	RF, UK, CN	Strong diplomatic presence; highly integrated (but diverse) IO / IW doctrines, strong industry links, heavily integrated w. DIME	Medium to high, but differ greatly in specialty (all three candidates)	Very extensive, but different MO (injects, SIGINT,..)	Medium-to-high, differing emphasis (int.sy vs. war etc.), some "deterrence"
Class 3 (V) "cyber power"	DE, FR, IL	Significant influencers (various focused), weaker IW doctrine but strong CND, some industry partner	Specialized, mostly "battlefield", some "tailored access"	Varies (Moderate-Extensive), most tailored	Hight to very high – strong centralized defensive structures
Class 4 (III-IV) "national cyber security)	Nordic, CH, J, NL, AU, ROK, ...	Specialist actors, usually defensive orientated, NCS doctrine paramount, some battlefield integration	Limited, some "tailored access" and MPAC-ish (mostly Pentest)	Limited, mostly SIGINT or tailored	Medium to very high – defensive orientation but readiness varies
Class 5 (II-III) "cyber aspirant"	IS, ID, PK, BR, IN, BR, ...	Developing actors, w. ambitions, largely "cybercrime" type activity	Some specialized, largely cybercrime – derived (e.g. DDoS)	Very Limited against hard targets (ok against soft)	Very Poor-medium Focus is on offensive

The Hague Centre
for Strategic Studies

HCSS
Lange Voorhout 1
2514 EA Hague

Follow us on social media:
@hcssnl

The Hague Centre for Strategic Studies
Email: info@hcss.nl
Website: www.hcss.nl