“One reason why deterrence has to be rescued from its current condition of semiretirement, is not so much because it offers great prospects of success, but rather because the leading alternatives suffer from severe limitations of their own” – Colin Gray (2003)

Introduction

In the analysis quoted above the military strategist Colin S. Gray investigates the rise and fall of deterrence theory from the heights of the Cold War - when it served as the core concept of international relations - to a situation in 2003 where it was on the verge of extinction. Instead, the US and others looked to strategies of preemption to confront new and apparently non-deterrable threats emanating from the Middle East and South Asia. With the Russian Federation’s use of “hybrid warfare” against Ukraine in 2014, and the following annexation of Crimea, deterrence has received renewed attention from scholars, military professionals and policy analysts alike as the “strategy of choice” in dealing with Russia’s revisionist challenge to the West. In Gray’s terms, deterrence is out of retirement.

The posture currently taken by NATO and the EU vis-à-vis Russia is fundamentally a two-pronged strategy of deterrence: by punishment at the conventional level, and deterrence by denial – through resilience – on what could reasonably be referred to as the “hybrid level”. This Information Note will focus on the latter, using Russian election interference as an empirical illustration. The basic argument is that whereas resilience – the “leading alternative” to paraphrase Colin Gray – is a necessary building block in creating a coherent strategy against Russia, it is nevertheless insufficient when it comes to deterring Russia from unwanted acts. To effectively enforce the rules of the game and change Russia’s behavior will require a renewed focus on deterrence by punishment below the threshold of war – on the hybrid level. Such a strategy would fundamentally serve to change the perception of the hybrid level as a zone of impunity, without provoking unnecessary escalation. In short, while the concept of deterrence might have been rescued in the wake of Russia’s resurgence, now is the time to renew and revitalise the strategy of deterrence.

The Response and its Shortcomings

The combined response by NATO and the EU to Russian revisionism in Ukraine and beyond can reasonably be understood within the framework of deterrence. In doing so, it is necessary to distinguish between deterrence initiatives primarily aimed at the conventional level as opposed to initiatives primarily addressing the Russian challenge on the hybrid level. It is also useful to inject the classic distinction between deterrence by punishment as opposed to deterrence by denial on each level.

This “return of deterrence” is clearly evident in the language of the NATO 2014 Summit. To cite two leading European scholars on the issue: “When NATO-allies met at their Wales Summit in September 2014, the D-word was back in vogue. Not in a muttering, shy or implicit way, but unambiguously and straightforward”. The stark rhetoric directed at Russia was repeated two years later in Warsaw where new initiatives – such as the Enhanced Forward Presence (EFP) deployment in the Baltic countries and Poland – aimed at deterring Russia from unwanted acts were introduced to supplement previous ones.

NATO’s initiatives seemed to be tailored with conventional scenarios in mind and premised on a deterrence by punishment logic. Yet initiatives on the conventional level also have implications on the hybrid level. On the one hand, Dr. Martin Zapfe has argued that the EFP deployment might bring with it new vulnerabilities for Russia to exploit through hybrid warfare. On the other hand, the proximity of NATO combat troops may reduce the propensity of Russia to use hybrid warfare in the Baltic theater in order to avoid the risk of moving up the escalation ladder. How these measures play out in reality remains to be seen. What is clear is that NATO and the EU address the hybrid level in its own right. This notion – NATO and EU’s strategic approach to the hybrid level – is the “centre of gravity” of this paper. The next section takes this further by showing how the deterrence logic applied on the hybrid level differs significantly from the one pursued at the conventional level.

Deterrence by Resilience as the New Black in the Gray Zone

The driving concept behind NATO and the EU’s approach to countering hybrid threats is essentially resilience. Understanding why resilience is the “go-to-strategy” requires an introduction to the problem itself, so this section starts by unpacking the concept of hybrid threats. It then moves on to discuss resilience as a response - as viewed by NATO and the EU – before examining the relationship between resilience and deterrence.

The term “hybrid threats” is shrouded in conceptual confusion. The reasons for this are (at least) two-fold. First of all, the concept of hybrid threats is often used interchangeably with hybrid warfare – in itself a contested concept – whereas in fact the two concepts are distinct and refer to different scenarios: the key aspect being whether or not the threat or use of force is included in the activity or not. Secondly, the term hybrid threats is used to refer to ongoing as well as future hostile activity, adding another layer of confusion to the concept.
Nonetheless, a growing consensus around the key characteristics of this challenge is emerging, based on four common elements. First of all, hybrid threats operate in the so-called “gray zone” between war and peace and are ambiguous by strategic design with an emphasis on non-kinetic instruments of power. Thus, hybrid threats deliberately challenge our ability to do timely early warning and detection given the fact that the abundance of the coercive actions against us take place in domains we traditionally consider outside the realm of war. Secondly, hybrid threats are characterized by an extensive use of proxies in domains where it is often challenging and time-consuming to ascribe responsibility for the actions taken. In short, this is the problem of attribution. Thirdly, hybrid threats challenge our ability to respond in the sense that while the actions might be malicious, they may also fall short of an outright armed attack. The problem related to response measures is exacerbated by the fluid and contested nature of the boundaries around lawful and legitimate actions. Fourthly, the fact that it is difficult to detect, attribute and identify a proper response against hybrid threats makes it a low-cost, low-risk strategy with potential for high gains. It follows, that we can reasonably expect to see hybrid threats as described here becoming a permanent phenomenon of international relations in the foreseeable future.

This understanding of hybrid threats has some conceptual commonality with the idea of “risk”. In particular, the amorphous nature of hybrid threats renders both their occurrence and their emergent effects essentially unpredictable. As V.M. Rasmussen says: “From a risk perspective a danger is much less computable [predictable] than from a threat perspective...the best one can hope for is to manage or pre-empt a risk”. Accordingly, as with risks the best approach to countering hybrid threats may be to manage, pre-empt or mitigate the likely consequences, rather than attempting to stop or prevent attacks in the first place. This is an adequate description of the “resilience” approach.

In trying to counter hybrid threats, NATO and the EU have introduced resilience as their preferred strategy. The 2016 Joint Declaration by the European Council, the European Commission and NATO – one of the key outcomes of the Warsaw Summit – diagnoses hybrid threats as a key challenge for the Euro-Atlantic community and prescribes large doses of resilience as the required medicine. The EU’s view of resilience is about improving the ability of its member states to absorb, adapt and recover from shocks, delivered through initiatives – including through enhancing democracy, human rights and the rule of law – both within the EU and in adjacent regions. NATO’s perception of resilience reiterates the idea of internal and external components, while making a more explicit connection between resilience and deterrence: “Military efforts to defend Alliance territory and populations [i.e. conventional deterrence] therefore need to be complemented by robust civil preparedness. However, civil capabilities can be vulnerable to disruption and attack in both peace and war. By reducing these vulnerabilities, NATO reduces the risk of a potential attack. A high level of resilience is therefore an essential aspect of credible deterrence”.

The link between resilience and deterrence brings us back to deterrence by denial. The deterrence logic is that the adversary may refrain from taking aggressive action if they regard their chances of success to be lower as a result of mitigating or preparatory measures taken by the actor targeted. In sum, resilience as currently understood and implemented by NATO and the EU is based on the capacity for absorption, adaptation and recovery in order to underpin a strategy of what could reasonably be labelled “21st Century deterrence by denial”.

The limitations of resilience

As with most strategy, an approach based on resilience has weaknesses. But before examining the limitations of resilience as a way of countering hybrid threats, it is important to note that resilience is indeed a necessary component in creating a robust strategy. Building physical, cognitive and legal resilience within societies, between states and on inter-organizational levels is a sound ambition. Yet it is also a challenging one, for two key reasons.

First, resilience is fundamentally a long-term project aimed at overcoming vulnerabilities across society that might not be easily amenable to change – and may even be an inherent part of the liberal-democratic model. The political and social landscape across the Euro-Atlantic region and beyond is characterised by varying degrees of polarisation. These forces challenge the very idea of comprehensive approaches to societal resilience, let alone the practicalities of implementation.

Second, even high levels of resilience are not a bulwark against hybrid attacks. An example is the trajectory of Russian interference in foreign elections since 2014; one study has found 18 cases of interference in this period, with 3 cases of substantial impact on the electoral result. In other words, electoral interference became established as a significant feature of international affairs despite efforts to increase societal resilience and democratic transparency. Yet while the strategy of resilience has seemingly not been able to deter Russia from these unwanted acts, the resilience measures in place are presumably a key component in mitigating the adverse effects of the attacks themselves.

The key question is where this trajectory leaves resilience as a strategy for countering hybrid threats. One conclusion might be that the negative effect of electoral interferences are in fact rather limited and that there is no need to exaggerate the issue at hand - from 18 is not a dazzling record of accomplishment. We might also reasonably assume the prospective marginal return of future electoral interference will diminish due to evolving resilience to these campaigns. Yet the fundamental problem with this line of reasoning is the acceptance of this kind of Russian behavior as a “new normal”. Viewed from this angle, the actions in themselves – and not their likely effects – are the key issue. Furthermore, there is ample reason to expect this pattern of behaviour to continue unabated in the years to come as electoral interference – and similar attacks on the hybrid level – becomes regarded as an effective, low-cost, low-risk strategy for revisionism. To paraphrase Susan Hennessey, the gray zone for Russia has largely become a “zone of impunity”. To change this situation, the cost-benefit calculus of revisionist actors such as Russia must be redressed. Yet it can be seen that resilience in itself is not going to provide the answer. While a necessary backdrop, deterrence by denial – through resilience – must be complemented by deterrence by punishment on the hybrid level. How to pursue this approach is explored in the next section.
Moving Beyond Resilience – The Case for Deterrence by Punishment on the Hybrid Level

The starting point for a new approach to deterrence by punishment on the hybrid level is to challenge some key assumptions about hybrid threats. First, while hybrid threats are often described as difficult to detect and attribute, it is critical to inject a distinction between difficult and impossible.21 Recent examples – such as interference in US and French elections, or the “Bundestag hack” in 2016 – show that it was in fact possible to trace and attribute responsibility. While detection and attribution will remain fundamentally troublesome, it appears that more often than not an accurate picture can be established to inform decision-making and public attribution. Furthermore, advances in AI and data-science may enhance veracity and reaction time.22

A second assumption which stands in the way of a new approach to deterrence on the hybrid level is that the options to respond to hybrid attacks are fundamentally limited. This perception exists for two main reasons: the “attribution problem” described above, and the lack of a readily available “playbook” of proven response options. In order to develop such a “playbook” and change the strategic calculus of aggressors who operate on the hybrid level, an obvious starting point to revitalize our current approach is the existing deterrence “syllabus”. This literature often argues that traditional deterrence theory does not help address emerging threats such as cyber and hybrid.23 But these perspectives are premature exactly because this kind of deterrence has not been done in a profound or convincing way. To do so requires returning to the classic deterrence “triad” of communication, capability and credibility.24

The first step towards a coherent deterrence by punishment posture is to identify and communicate thresholds of response. In other words: what actions are deemed unacceptable? Establishing the limits of unacceptable behavior is a necessary task in the sense that one should not strive to deter malicious activities per se – this would be a Herculean effort, and is exactly where resilience must “do the job”. Instead, deterrence by punishment efforts should be directed against the most severe hybrid threats while acknowledging that introducing such a threshold may also be viewed as an “invitation to act” below it. In sum, it must be made public to our adversaries that some actions are beyond the pale and will be met with punishment measures.

The next step is regarding the capability to inflict pain on the adversary. The objective here is to create a situation where the adversary perceives that the costs likely to be incurred through hostile activity will outweigh potential gains. To achieve this effect the punishment measures must be forceful and calibrated towards key vulnerabilities of the opponent. The guiding principle must be: What does the potential aggressor not want to happen? In the case of Russia, one example would be the so-called Magnitsky Act, put to the table in the bilateral presidential meeting in Helsinki in July 2018.25 However, a wider view of Russia’s vulnerabilities goes well beyond the economic sphere.26

Another important aspect of the capability to respond to hybrid threats is the basis of any response in international law. Although the “legal edifice” is not entirely constructed, when it comes to countering hybrid threats there is in fact a legal basis for responding to and punishing hostile, aggressive and illegal activity such as electoral interference and wrongful territorial claims. In short, a legal basis for responding to hybrid attacks can be established if the measures taken are proportional and aimed at bringing the offending actor back in line, or because of manifest lack of due diligence by states hosting non-state actors engaged in malign activities.27 Importantly, the countermeasures can be done either “in-domain” or “outside-domain”.28

This brings us to the final issue concerning credibility. The aggressor must believe they will incur a specific punishment in the case of a transgression. Their assessment will presumably depend on various factors – not least, the question of who their opponent is. Although it might be possible to muster a credible deterrence policy for select individual states with sufficient resources, deterrence remains a collaborative endeavor for the vast majority of states.

Returning to the case of Russia, this is exactly where NATO and the EU, as the centrepieces of the Western security architecture, enter the equation. The basic strategy of deterrence by denial through resilience measures, as described above, requires recalibrating if a situation is to be avoided where hybrid aggression towards their member states becomes the norm. This strategy needs to be recalibrated exactly because resilience on the one hand side is too little and the threat of escalation to Article 5 is too much in the sense of being a credible response to low-level hybrid activity. The problem with malicious actions at the hybrid level is that while they are significant intrusions that can have profound effects, they are not considered escalation material as such. This “loophole” is of course well known to proponents of hybrid aggression. Therefore, the guiding principle for a revitalized deterrence by punishment strategy would be to communicate our threshold and to identify credible punitive actions tailored towards key vulnerabilities of the adversary while staying below the threshold of an armed attack. That such punitive measures are in accordance with international law, and therefore possible to legitimate in public underlines the credibility of the threat. In doing this, both NATO and the EU need to deploy their levers of power in a synchronized fashion.

Doing “deterrence by coalition” is of course not an easy job considering different threat perceptions, risk profiles and the like. This issue is of course not new to neither NATO nor the EU individually and a well-known challenge to cooperation between multinational institutions. However, these “cohesion-obstacles” are not insurmountable. One example is how the EU has managed to sustain “middle range” punitive actions – such as the sanctions regime directed at Russia – despite such challenges to threat perception amongst member states. In the same vein, Russia’s resurgence has revitalized NATO-EU partnership and made it clear to both organizations that they are interdependent more than ever and in combination possess a strategic toolbox of a considerable magnitude.29

The Flipside of Deterrence by Punishment – the Fear of Escalation

One of the greatest reservations against communicating a willingness to take punitive actions concerns the risk of escalation and increased tension due to a more assertive or offensive posture. Yet when looking into possible responses – retaliation in cyberspace, for example – a number of “self-dampening”
mechanisms appear to be in place that may be applicable to many types of responses. One example is the requirement to establish some level of attribution of aggression on which to base a response. As discussed above, while attribution is rarely impossible, it can be a time-consuming and technically-challenging endeavor. In a high-stakes scenario the time taken to get attribution as right as possible means there will be plenty of time to think twice about actions and consequences, and to lean on diplomatic measures in parallel. An example of a self-dampening mechanism related to the cyber domain is the large investment required to develop credible offensive capabilities. Moreover, an offensive cyber capability is a transitory tool. The ability to access a computer system or network to cause harm or damage is only temporary and dependent on a very rapid and ongoing patching of vulnerabilities. At the same time, wielding the cyber instrument despite the downside of “burning” the capability might have a de-escalatory effect by communicating capability and credibility to the opponent with a view to discouraging future hostile attacks. Targeting in any domain – including cyberspace – must also follow relevant rules, law and due-process which will self-limit the range of targets and actions available.

A final “escalatory showdown” is related to the challenge of identifying and developing targets of adequate strategic significance – not too much, not too little – to achieve the desired effect. Responding to hybrid aggression by applying “middle range” punitive actions that are proportionate to the aggression threatened or suffered will also self-limit the escalatory potential. It is unlikely to be in the interest of any hybrid aggressor to pursue an escalatory spiral above and beyond where they were looking to compete in the first place: on the hybrid level. Nevertheless, one way to mitigate the risk of escalation – while enhancing civilian oversight and interagency coordination – would be to establish rules of engagement for punitive actions on the hybrid level. This would provide decision-makers with common guidelines to pursue punitive actions that fall below the “use of force threshold”. Moreover, this could actually bolster the credibility of punitive actions by signaling to hybrid aggressors the intent to take pre-prepared punitive actions when deemed necessary: in other words, a “playbook” for countering hybrid threats.

**Time to think creatively**

The purpose of this Information Note has been to cast light upon the inadequacies of much current thinking and strategy for countering hybrid threats, with a particular empirical focus on the case of Russia, the EU and NATO. It has been shown that whereas resilience is a sound and necessary building block in countering hybrid aggression, it is nevertheless insufficient if the ambition is to alter the cost-benefit calculus of the adversary and change their behaviour. Instead, NATO and the EU need to think in terms of **deterrence by punishment on the hybrid level** in order to arrive at a more coherent and robust strategy. In doing so this paper has also shown that it is in fact possible to overcome many of the obstacles traditionally thought to make hybrid threats “undeterrable”. Clearly, this proposed departure from the existing framework will entail challenges and associated risks. However, the time has come for the EU and NATO to think creatively about their responsibility and the opportunities to muster an effective strategy to counter hybrid threats to change the perception of the gray zone as a “zone of impunity” and regain the initiative against hybrid aggressors.

---

**Endnotes**

1. This Information Note is the original and extended version of “Going Beyond Resilience – A Revitalized Approach to Countering Hybrid Threats”, European Center of Excellence for Countering Hybrid Threats, 2019, available at: https://www.hybridcoe.fi/publications/strategic-analysis-2018-going-beyond-resilience-a-revitalised-approach-to-countering-hybrid-threats/


3. The authors’ perspective on current Russian ‘revisionism’ is on a par with the idea of “measured revisionism” by Dr. Michael J. Mazarr in the monograph Mastering the Grey Zone – Understanding a Changing Era of Conflict, pp. 24-25. Russia’s selfperception as a great power with special responsibilities, actively striving for a polycentric world order, is e.g. made explicit in the current National Security Strategy from 2015 and The Foreign Policy Concept of the Russian Federation from 2016 (available at: http://www.mid.ru/en/foreign_policy_oficial_documents/-asset_publisher/CptICkB6BZ29/content/id/2542248?p_id=101_INSTANCE_CptICkB6BZ29&101_INSTANCE_CptICkB6BZ29_languageId=en_GB). However, the idea of restoring Russia as a respected great power on the world stage has been a persistent theme in formal strategy documents and public speeches throughout the Putin-era and has a much longer trajectory in Russian strategic thinking and philosophy as well. See for example: Neumann, I. (2008): Russia as a Great Power – 1815-2007, Journal of International Relations and Development.

4. It is necessary to stress, that the demarcation between a conventional and hybrid level does not in any way imply that the levels should be understood as isolated realms. On the contrary, the levels are very much interdependent. Thus, strategies and actions on e.g. the conventional level will almost by default have implications on the hybrid level and vice versa.


7. As Martin Zapfe states: “The main function of NATO’s EFP is to help deter a conventional Russian attack by providing a tripwire, the engagement of which would all but guarantee that the Alliance as a whole would respond in some way”: Zapfe, Martin (2017): Deterrence from the Ground up: Understanding NATO’s Enhanced Forward Presence, p. 150, Survival, vol. 59, no. 3, June-July 2017, pp. 147-160.

8. Whether the initiatives from Wales and Warsaw combined will in fact “do the job” and deter Russia from conventional attacks has been a matter of intense debate ever since - an assessment of which is unfortunately beyond the scope of this paper.

9. Ibid.


14. Joint Declaration by the President of the European Council, The President of the European Commission, and the Secretary General of the North
Deterrence by Punishment


18 It is worth mentioning, that NATO reserves the right to treat a cyber- or a hybrid attack as equivalent to an armed attack with reference to Article 5 of the Washington Treaty. Translated into “deterrence language” this could rightly be viewed as a case of deterrence by punishment. The view presented here is, however, that the option is reserved for clear-cut cases only. Thus, both state and non-state actors appear to have sufficient leeway for conducting malicious activities on the hybrid level without necessarily having to fear a conventional response in return.


