Integrated Deterrence: NATO’s ‘First Reset’ Strategy

GLOBSEC
NATO
ADAPTATION
INITIATIVE
Supporting Paper
The GLOBSEC NATO Adaptation Initiative, led by **General (Retd) John R. Allen**, is GLOBSEC’s foremost contribution to debates about the future of the Alliance. Given the substantial changes within the global security environment, GLOBSEC has undertaken a year-long project, following its annual Spring conference and the July NATO Summit in Warsaw, to explore challenges faced by the Alliance in adapting to a very different strategic environment than that of any time since the end of the Cold War. The Initiative envisages a series of policy papers which will address the nature of NATO adaptation and the challenges it must overcome if it is to remain a viable and credible alliance for the peace and stability in the transatlantic area. The policy papers published within the GLOBSEC NATO Adaptation Initiative are authored by the Initiative’s Steering Committee members: **General (Retd) John R. Allen, Admiral (Retd) Giampaolo di Paola, General (Retd) Wolf Langheld, Professor Julian Lindley-French, Ambassador Tomáš Valášek, Ambassador Alexander Vershbow** and other acclaimed authorities from the field of global security and strategy. The aim of the involvement of such a wide array of experts is to reinforce the unique partnership between policy-makers, military leaders and leading academics and commentators. These policy papers will prelude and result with the publication of the Initiative’s Steering Committee Recommendation Two Pager and the Main Report to be launched in October 2017. The Interim Report will be released during the GLOBSEC 2017 Bratislava Forum.

These outputs will be augmented by shorter policy papers (on cybersecurity, A2/AD capability, intelligence, and threats emanating from the South) prepared by the GLOBSEC Policy Institute between January and October 2017.

Integrated Deterrence:
NATO’s ‘First Reset’ Strategy
GLOBSEC NATO Adaptation Initiative
Paul Cornish

Introduction

Deterrence is no mystery; it is a feature of many human activities, behaviours and relationships, ranging from the private matter of bringing up children, to society’s attempts to control crime. At any level, and in any sector, deterrence is a promise to impose a cost on a given action in order that the potential perpetrator is convinced that any perceived benefits of the action will be outweighed by the costs incurred, and will thus choose not to act as planned or threatened.

Several ingredients must be in place for deterrence to function, often described as the ‘three Cs’. First, the deterrer must have the capability to impose the costs he has promised or threatened. Second, the deterrer’s promise must seem credible to the potential miscreant. As well as the appropriate capability, credible deterrence also requires that the deterrer has the will – personal, political or moral – to carry out her promise, and that this can be communicated to, and understood by, the wrong-doer. Deterrence is therefore a relational activity, in which both sides must employ a broadly compatible rationality.

Deterrence is particularly well known as a feature of politico-military strategy. The basic ingredients remain: a potential aggressor’s cost-benefit calculation might be influenced by the threat of a punitive response, or by the realisation that the defender’s preparations are so advanced and effective that the costs of carrying out the aggression would be too great, whether politically, financially, militarily or even reputationally. Throughout human history, when an aggressor has taken stock and decided not to proceed, it is possible – but not certain – that deterrence will have played a part in that decision. This uncertainty is discussed briefly below.

Yet politico-military strategic deterrence is far from what it was. The practical and intellectual underpinnings of this crucial idea have been fading fast from the popular and political memory – just when the need for deterrence could scarcely be greater. Alarmingly, this damning judgement might even be said of NATO, a politico-military alliance whose raison d’être was – and must continue to be – deterrence. If deterrence is failing, then the same must be said of politico-military strategy – including that offered by the Alliance.

The intention of this paper is, first, to ask how it is that we have arrived at this point; how it is that the logic of deterrence has declined so markedly in strategic and popular culture. The aim then is to describe a way out of this dilemma. The paper begins with an account of the development of deterrence during the Cold War, before describing the ‘crisis of scepticism’ that has beset deterrence thinking and practice in the 21st century. This wave of scepticism might not be significant, were it not for the fact that Russia, having detected the loss of faith among NATO’s nations, has grasped the opportunity to challenge NATO’s strategic deterrence in the manner that no deterrence posture should be expected to tolerate. Russia is calling NATO’s bluff.

This paper makes three main arguments. First, NATO should work urgently to revive its deterrence posture; a posture which should be integrated vertically, horizontally and functionally. Second, a ‘future-proofed’ deterrence posture should be a central component of the programme of adaptation set in motion at NATO’s 2016 Warsaw Summit. Third, the paper argues that a revived and integrated deterrence posture should have no other purpose than deterrence itself. The revival of NATO

1 Chief Strategist, Cityform Public Policy Analysis Ltd., London and Professorial Fellow, National Security College, Australian National University. I am grateful to General (Ret.) John R. Allen and Professor Julian Lindley-French for their comments on an earlier version of this paper.
deterrence is not an opportunity to compensate for some defence deficiency or another, under the guise of an ‘Offset Strategy’ of some sort. Instead, this is the moment for a ‘First Reset Strategy’ – a co-ordinated effort to rediscover and then achieve the fundamental goal of deterrence: the establishment of order, even in an adversarial environment. The paper concludes by recommending five concrete actions to be taken as soon as possible by NATO in order to launch the First Reset Strategy.

**Cold War Deterrence**

With the invention of atomic and nuclear weapons in the mid-20th century, and the beginning of the Cold War, it became critically important that deterrence should work – always. But the evolution of nuclear deterrence was not quite as automatic as might be supposed. After their use against Japan in August 1945, there was for some years a tendency to see atomic weapons as ‘super-bombs’, and as a means to extend and amplify existing doctrines of strategic air power. Atomic weapons also offered more ‘bang for the buck’ than an expensive conventional force posture and could offset weaknesses in conventional defences, particularly at a time when the conventional strength of the Soviet Union was thought to have remained overwhelming while the US and its European allies had demobilised rapidly after the war. The notion that one technology or another can ‘offset’, or compensate for, conventional weakness is still alive in the security and strategy debate today, in the form of the Third Offset Strategy.

An offset strategy is the use of technological superiority to compensate for perceived imbalances and weaknesses in conventional military strength. In a sense, the ‘offset’ is an attempt to re-establish the capability, credibility and communications upon which any successful deterrence posture must be based. The first such strategy was developed in the early 1950s, in the form of Eisenhower’s New Look Strategy whereby increases in nuclear capability (including at the tactical and operational levels) would offset Soviet and Warsaw Pact conventional military advantages. The Second Offset Strategy was a product of the 1970s and 1980s when doctrinal and technological developments such as ‘Follow-On Forces Attack’ and ‘Airland Battle 2000’ would enable non-nuclear attack against Soviet and Warsaw Pact echelons deep in their own territory. Developed under the Obama Administration, the goal of the Third Offset Strategy is to acquire ‘the means to offset advantages or advances in anti-access area denial weapons and other advanced technologies that we see proliferating around the world.’

As the Cold War advanced, atomic and then nuclear deterrence became far more elaborate. Capability alone was not sufficient; with weapons of this scale, credibility and will also mattered. And as the vulnerability to attack became mutual, so communication became ever more important. Atomic and nuclear deterrence could no longer be considered a component of one side’s politico-military strategy; mutual deterrence became the central purpose of the Cold War as a whole. Paradoxically, this critically important idea could never be analysed too closely, for fear of revealing its fragility; it could not fail, but neither could it be tested. In his seminal *Strategy in the Missile Age* (1959), Bernard Brodie pointed to the difficulties of deterrence in the nuclear era; ‘We expect the system to be always ready to spring while going permanently unused.’

Perhaps inaction speaks louder than words. It is often claimed that for four decades after its creation in 1949 NATO’s deterrent posture ‘worked’, since the Cold War in Europe never became ‘hot’. The Soviet Union and its allies in, what was intended to be the counterpart to NATO, the Warsaw Treaty Organisation (or ‘Warsaw Pact’ as it was commonly termed) never attacked the NATO treaty area. The difficulty with this claim, and with deterrence thinking generally, is the problem of negative proof. It will always be difficult, practically and logically, to isolate the reasons why aggression or war did not take place and equally difficult, therefore, to be confident that deterrence had succeeded as the cause, so to speak, of a non-event. Conversely, the frequency and extent of war in human history suggest that it might be easier to find evidence of deterrence having failed, or not having been attempted.

The end of the Cold War has been described as the ‘bonfire of the certainties’ – a particularly apt expression when it comes to the consideration of deterrence in Europe. A body of ideas, capabilities and untestable assumptions that for several decades had nevertheless been at the heart of security policy in and for Europe, deterrence was unceremoniously consigned to history in the early 1990s. It was not only that the highly elaborate, mutual strategic deterrence of the Cold War was considered suddenly to have become irrelevant but that politico-military deterrence itself was thought to be obsolete. In retrospect, this judgement was both complacent and ahistoric. The past 25 years have shown that in certain respects, mutual strategic deterrence is still required; and there should have been no reason to suppose that the basics of deterrence had become any less valid with the end of the Cold War than they had been throughout the previous millennia of human history. Nevertheless, almost overnight European deterrence arguably became a mere façade, swaying ominously in even the lightest of strategic breezes with no clear sense of how and where it is anchored.

Former Chairman of the US Joint Chiefs of Staff General Martin Dempsey observed succinctly that ‘As an alliance [i.e. NATO] we’ve taken deterrence for granted for 20 years now, but we can’t do that anymore.’ Deterrence is in a crisis of scepticism; a situation that should be a cause of concern in NATO, for NATO governments and for European security in general.

21st Century Deterrence: The Crisis of Scepticism

Since the signature of the Washington Treaty in 1949, the core task of the North Atlantic Treaty Organisation has been to defend the territory and interests of the treaty signatories in North America, in Europe and in the north Atlantic Ocean. In spite of any uncertainty prompted by the problem of ‘negative proof’ discussed earlier, it seems likely, at the very least, that deterrence forms part of the explanation for the stability of the NATO area between 1949 and the end of the Cold War some forty years later. But if it is reasonable to conclude that deterrence did what was expected of it during the Cold War, why is it now in a state of crisis?

Part of the explanation is that the European security order is challenged in so many ways and on so many levels that Cold War style deterrence thinking now appears scarcely relevant; little more than an historical artefact. On the one hand, the shibboleths of Cold War deterrence seem too narrow and monolithic, too concerned with a very specific, militarised conception of order and security. On the other hand, these ideas are seen by some to run the risk of becoming a self-fulfilling prophecy. Advocates of deterrence use the language of order and stability. But in practice, runs the argument, deterrence is manifestly a preparation for everything but order and stability. For those western European polities that emerged from the Cold War in the early 1990s, the so-called ‘peace dividend’ was a far preferable alternative to the continuation of Cold War levels of defence spending and preparations for conventional war in Europe.

Another part of the explanation lies in the form of a category error; a confusion between underlying principle and contingent effect. Carl von Clausewitz, a nineteenth century Prussian general and a philosopher of war, suggested that ‘war can be thought of in two different ways – its absolute form or one of the variant forms that it actually takes.’ Current scholarship tends to interpret this to mean that, for Clausewitz, war had both an unchanging nature and an ever-changing character. Taking a broadly Clausewitzian line it could be said that deterrence, like war, also has an unchanging nature and that it too can take on as many forms as history and circumstance demand. It seems all-too easy, however, to confuse the enduring with the contingent. In the UK and elsewhere some military practitioners and analysts came to the mistaken view that desert warfare in Afghanistan and Iraq revealed the fundamental nature of modern war rather than serving as a mere illustration of its ever-

3 ‘Bonfire of the certainties’ – an expression coined by George Robertson (now Lord Robertson) in a speech at Chatham House in 1990.
5 By one account, Clausewitz was the philosopher of war. In the late 1970s the English philosopher Philip Windsor described Clausewitz’s On War as ‘the only work of philosophic stature to have been written about war in the modern period.’ P. Windsor, ‘The Clock, the Context and Clausewitz. Millennium (Vol. 6, No. 2, 1977), p.193.
changing character. Similarly, it would be a mistake to confuse one, contingent form of deterrence with its unchanging nature and then to assume that form must therefore be universally applicable and unfailingly durable. Thus, any expectation that Cold War deterrence thinking and practice could be transferred unaltered to post-Cold War circumstances was always doomed to end in disappointment. After all, how straightforward and productive could it ever be to take one set of highly context-specific deterrence ideas and principles and transfer them, entire, to a completely different context – particularly when that new context is still evolving? But the consequences of this category error could scarcely be more serious for geostrategic stability and security. Because the Cold War forms of deterrence have been found wanting in the current strategic environment, so it is assumed that the core principles of deterrence have nothing to offer; the all-important ‘baby’ has been thrown away with its cold and unwanted ‘bathwater’.

Loss of interest in the principles of deterrence is happening at a time when the need to understand and then to operationalise these principles could scarcely be more pressing. It is for this reason that the current moment is described as a crisis. The stability and security of Europe requires a rediscovery and recategorisation of deterrence to meet a broad range of challenges, both traditional and novel. Some of the questions to be asked of deterrence are familiar, others much less so. But it is precisely this combination of the old with the new which poses a general challenge to European deterrence theory, policy and practice. In early 21st century Europe, what forms should deterrence take and how should it be achieved? Is mutual nuclear deterrence credible after the Cold War? How much conventional military deterrence is needed? How can terrorists and criminals be deterred? Is it possible to deter or prevent the use of chemical, biological, radiological and nuclear weapons by non-state actors? Can cyber deterrence be a meaningful activity? And can deterrence be achieved through other, non-military means such as sanctions and soft power?

There have been occasional attempts to revisit, revise and update the deterrence debate, but these have so far made little impression on a post-Cold War European security consensus which has proved to be complacent and lacking in vitality where deterrence is concerned. In 2003 the European Union published a European Security Strategy in which the words ‘deterrent’ and ‘deterrence’ did not appear.1 The 2016 European Union Global Strategy makes a slightly better effort: ‘As Europeans we must take greater responsibility for our security. We must be ready and able to deter, respond to, and protect ourselves against external threats.’2 Although this sentence is printed in a prominent, eye-catching way in the published document, it would be difficult to see in this form of words anything more than a passing acquaintance with a body of complex and important ideas.

In the UK’s case, although the National Security Strategy and Strategic Defence and Security Review (NSS & SDSR) of 2015 dwelt at some length on the need for a comprehensive, cross-governmental approach to deterrence, the document fell short of providing a publicly accessible understanding of 21st century deterrence; a conceptual framework in which policy, strategy, ethics, technology and psychology are all intertwined. Instead, the UK National Strategy took little more than a declaratory line, using language that offered no more than a basic understanding of the topic: ‘Deterrence means that any potential aggressors know that any benefits they may seek to gain by attacking the UK will be outweighed by the consequences for them.’3 Similarly, the most recent French White Paper on Defence and National Security (the Livre Blanc, published in 2013) also covered deterrence to some depth, but almost exclusively in terms of the preservation of France’s national nuclear deterrent capability.4

Deterrence faces a crisis of scepticism as to its relevance and its strategic significance. The cause of this scepticism is, in part, the Cold War-weariness that set in across Europe in the early 1990s and by the desire to spend public money in more immediately gratifying ways than on military personnel and

Calling NATO’s Bluff: Russia’s Strategy of Disruption

In late February 2014, in the aftermath of the Ukrainian revolution, pro-Russian protests in Crimea led to Russian forces and their proxies occupying strategically important areas of the Crimean Peninsula. Crimea had been transferred to the Ukrainian Soviet Socialist Republic by Khrushchev in 1954 and subsequently became an Autonomous Republic of independent Ukraine in 1991. Crimea was, in other words, an internationally recognised part of Ukrainian territory. In March 2014, however, a referendum (held while under conditions of military occupation) provided the questionable rationale for Russia’s annexation of Crimea. The annexation was, and remains, in breach of international law, but that detail mattered little in Moscow. Pro-Russian demonstrations in the Donbass region of Eastern Ukraine then prompted further Russian intervention. By late 2014 several thousand Russian troops were deployed inside Ukraine – an internationally recognised, independent, sovereign state – with many more poised on the Russian side of the border.

The challenge to the European order was stark: contrary to decades of expectation, European borders had once again been changed by armed force. The challenge to NATO was especially vivid. Although the West had, arguably, misunderstood and mishandled Russia since the end of the Cold War, and although Ukraine was not a member of the Alliance, Russia’s readiness to use armed force against its neighbour called NATO’s credibility into question. NATO had reinvented itself since the end of the Cold War as the sponsor of strategic stability in Europe, working with former adversaries. This claim no longer carried much weight.

What also impressed NATO’s strategic analysts and decision-makers was the manner in which Russia used coercive force in Crimea and Donbass. Terms such as ‘asymmetric’, ‘hybrid’ and ‘ambiguous’ warfare have for some years captured the imagination of strategic analysts. But Russia appeared to go one or two steps further in its use of so-called ‘new generation warfare’ – a combination of civil action and intimidation with the deniable use of conventional military forces, all orchestrated by a strategic communication campaign drawing heavily on the high-minded rhetoric of democratisation and humanitarian intervention. All in all, Russian actions in Crimea and Eastern Ukraine appeared to leave NATO with little, if anything to say or do; the Alliance’s bluff seemed to have been called.

Although some saw in these events the beginning of a ‘new Cold War’, NATO’s response was very much more cautious than might have been seen during the Cold War. Alliance air forces flew surveillance missions to monitor Russian military activity and an ‘air policing’ operation was launched over Estonia, Latvia and Lithuania. Naval vessels were deployed to the Baltic and the Mediterranean and military exercises took place in northern and central Europe. Yet in spite of these various deployments, decisions and initiatives, the Alliance seemed to have been wrong-footed by Russia. Russian intentions towards Ukraine continued to dominate NATO’s geostrategic analysis during 2015. Following the failure of the September 2014 Minsk Protocol, in February 2015 Russian President Vladimir Putin met his Ukrainian


11 Asymmetric warfare is best understood as war between one side which is militarily weak yet determined and ingenious, and another side which is militarily powerful yet complacent and inattentive. See Paul Cornish and Andrew Dorman, ‘Complex security and strategic latency: the UK Strategic Defence and Security Review 2015’, International Affairs (Vol. 91, No.2, 2015), p.357. The scope of hybrid warfare has been defined as follows: ‘Any adversary that simultaneously and adaptively employs a fused mix of conventional weapons, irregular tactics, terrorism and criminal behavior in the battle space to obtain their political objectives’: Frank G. Hoffman, ‘Hybrid vs. compound war. The Janus choice: defining today’s multifaceted conflict’, Armed Forces Journal, October 2009, http://www.armedforcesjournal.com/hybridvs-compound-war/.

12 ‘New generation’ (or ‘permanent’) warfare is widely considered to be a Russian innovation. It has been defined as follows: ‘The Russian view of modern warfare is based on the idea that the main battle-space is the mind and, as a result, new-generation wars are to be dominated by information and psychological warfare, in order to achieve superiority in troops and weapons control, morally and psychologically depressing the enemy’s armed forces personnel and civil population. The main objective is to reduce the necessity for deploying hard military power to the minimum necessary, making the opponent’s military and civil population support the attacker to the detriment of their own government and country’: Jānis Bērziņš, Russia’s new generation warfare in Ukraine: implications for Latvian defense policy (Riga: National Defence Academy of Latvia, Policy Paper No. 2, April 2014), p. 5.

7
counterpart, Petro Poroshenko in yet another attempt to resolve conflict in the Donbass region. The outcome, known as Minsk II, was an elaborate agreement which, by the end of 2015 had largely failed to achieve its goals. It was significant that NATO had little or no involvement in a strategic ceasefire agreement being negotiated on the edge of its treaty area. But the Ukraine crisis prompted rather more fundamental – even existential – questions about the Alliance. What was NATO actually for? What could NATO do well and effectively as a strategic politico-military alliance? How should NATO be configured to meet European and international security challenges? And did NATO, somehow, become ‘self-deterred’ in the face of Russian intervention?

For some politicians and strategic analysts, the continuation of Russia’s aggressive and unpredictable behaviour along its borders was ample proof that NATO had miscalculated as to the nature of the strategic threat to Europe. Terrorism and organised crime would, of course, continue to be significant security challenges but the real threat had now become clear; territorial aggression within Europe. Since the collapse of the Warsaw Pact and then the Soviet Union in 1991, NATO leaders had assumed, not unreasonably, that the Alliance’s long-standing commitment to territorial self-defence in western Europe need no longer be the focus of defence planning and expenditure. Perceptions of Russian adventurism, however, in Ukraine and elsewhere, prompted anxious questions as to whether NATO should seek to re-acquaint itself with Cold War-style conventional warfare in Europe, perhaps even to the extent of reinvesting in heavy armour formations (battle tanks, tracked infantry fighting vehicles, self-propelled artillery and so on). Was NATO’s purpose therefore to provide for defence (in the traditional sense), or for security (in the post-Cold War sense)? Or perhaps both?

NATO’s summit meeting in Warsaw from 8-9 July 2016 was an opportunity for the Alliance to show that its 28 member nations not only understood the gravity of the situation in which they found themselves but that they could also do something meaningful about it. As so often in NATO’s recent past, however, the first of these tasks proved easier to manage than the second. Earlier in the year NATO’s Secretary General Jens Stoltenberg had warned against ‘a more assertive Russia. Intimidating its neighbours. And changing borders by force.’\footnote{Joint Press Conference with NATO Secretary General Jens Stoltenberg and Antoni Macierewicz, Minister of Defence of Poland, 31 May 2016: http://www.nato.int/cps/en/natohq/opinions_131715.htm.} The language of the Warsaw communiqué was similarly stark. The document echoed Stoltenberg’s words, referring repeatedly to Russia’s ‘aggressive actions.’\footnote{North Atlantic Treaty Organisation, Warsaw Summit Communiqué, Press Release (2016) 100, 9 July 2016, paras. 5, 16, 19, 118: http://www.nato.int/cps/en/natohq/official_texts_133169.htm.} Although the Russian Question loomed very large at Warsaw, the communiqué took a wider view, describing an ‘arc of insecurity and instability along NATO’s periphery and beyond’ and noting that turbulence in the Middle East and North Africa could have ‘direct implications for the security of NATO.’\footnote{Ibid., paras. 5, 25.} But the signatories of the communiqué also sought to convey a sense of mature self-confidence. Approaching the 70th anniversary of its founding treaty, this was a highly experienced politico-military organisation wishing to appear firm and decisive yet at the same time calm, non-confrontational and, above all, in control of its strategic environment. NATO leaders knew their script well enough; just as during the decades of the Cold War, adversaries and aggressors must be met with a convincing deterrent posture and an equally convincing defensive structure. And just as deeply embedded in the Alliance’s strategic culture is the importance of talking to adversaries, particularly when times are tense. The Warsaw summit therefore centred around three familiar themes: deterrence, defence and dialogue.

These three themes are closely intertwined - strategic dialogue is conducted not only through diplomatic channels but also through military preparations and deployments. At their previous summit in Wales in September 2014, NATO leaders had agreed a Readiness Action Plan and established a Very High Readiness Joint Task Force.\footnote{North Atlantic Treaty Organisation, Wales Summit Declaration, Press Release (2014) 120, 5 September 2014: http://www.nato.int/cps/en/natohq/official_texts_112964.htm.} The Warsaw summit then tried to send even clearer signals with the decision to establish a ‘forward presence in the eastern part of the Alliance’;\footnote{Warsaw Summit Communiqué, para. 11.} four multinational battlegroups, headed by the UK, Canada, Germany and United States would be deployed to Estonia, Latvia, Lithuania and Poland respectively.\footnote{Ibid., para. 40.} Importantly, this commitment – described in Warsaw as a...
‘robust multinational presence’ – would be open-ended. Yet it was nevertheless a very small military commitment. To those unfamiliar with military organisations, the expression ‘battlegroup’ might suggest a large, highly capable formation. But the battlegroup is a tactical unit, typically involving 600-1,000 infantry and armoured troops as well as combat support elements such as reconnaissance, artillery, engineers and logistics. The battlegroup is not, therefore, a large, operationally capable formation and is not usually considered sufficient for independent deployment. Using language coined during the Cold War, some have described these deployments as a ‘trip-wire’. But a trip-wire only functions as such if it is being watched closely and if something decisive (and usually explosive) happens when it is tripped. In the absence of very significant reinforcements – equipped, trained and poised to react when needed – these battalion-based deployments could not qualify as a military trip-wire; the most that could be said of them is that they might serve as a ‘political trip-wire’ of some sort.

In order to make this limited forward presence appear more meaningful as a military trip-wire, towards the end of 2016 unnamed ‘NATO commanders’ were reported to be arguing for the response time of the Alliance’s ‘follow-on-force’ of c.300,000 troops to be reduced from six months to about two months. If achievable, this change in response time would be a marked improvement upon previous, rather tentative and unconvincing assertions about NATO force readiness, and the strategic significance of such a step-change would not be lost on Russia. But will the follow-on force always be available in these numbers and will a reaction time of two months be fast enough in any case? For Estonia, for example, the deployment of a NATO battlegroup in Spring/Summer 2017 will almost certainly be welcomed as a symbol of Alliance cohesion, but probably not much more; in the two months that NATO would need to come to Estonia’s defence, Russian forces could very rapidly and easily have achieved any geopolitical goals they might have. According to an assessment published in 2016 by senior research analysts at the RAND Corporation, Russian ground forces in the Western Military District (adjacent to the Baltic Republics) comprise as many as 22 manoeuvre battalions, ten artillery battalions, four missile battalions and six attack helicopter battalions, with further naval infantry and missile units available on reserve in Kaliningrad. The authors of the RAND study estimate that these forces could reach the outskirts of the capitals of Estonia and Latvia in just 60 hours.

The unavoidable conclusion to be drawn from this case study of the Russian strategic challenge to NATO is that having at least reminded itself of the idea of deterrence at the 2016 Warsaw summit, the Alliance nevertheless still has some way to go if it wishes to bring the idea back to life. Military deployments are, certainly, an important method of communication to an actual or potential adversary. But communication alone is not sufficient; successful deterrence requires the ‘three Cs’ – the communication of politically credible strategic capability. NATO must – and can – improve the quality and coherence of its response.

NATO’s Response: Integrated Deterrence

Contemporary challenges to the European security order require an intelligent reapplication of principles and techniques which have long been familiar. The basic principles of deterrence are unchanging; the imposition of intolerable cost, whether through defence and denial or by the threat of a punitive response, to dissuade an actual or potential adversary from adventurism or aggression. The ‘three Cs’ remain a valid description of the conditions that a deterrent posture – any deterrent posture – must satisfy if it is to be effective and if an adversary’s cost-benefit analysis is therefore to be influenced appropriately. But effective, durable deterrence has become more than strong defences and much more than a declared promise to inflict violence with one capability or another; 21st century


European deterrence requires a range of capabilities (passive and active, military and non-military) in a posture which is both coherent and credible and which can be communicated unambiguously to any potential adversary.

To borrow Clausewitz’s distinction once again, if its nature is unchanging then what is the character of 21st century strategic deterrence? Essentially it is that deterrence has not one character, or form, but several. NATO’s deterrent challenge, therefore, is threefold. First, NATO must identify where and how deterrence is needed. What are the current and likely challenges to NATO’s territory and interests, from what source and at what level? What do potential adversaries have at stake; how much risk are they willing to bear in mounting those challenges? Second, the governments of NATO nations must ensure that its deterrent needs, having been identified, are then properly funded and turned into capability. Third and finally, NATO and its member governments must decide how to integrate these different deterrence capabilities into one coherent whole. Without an overarching, unifying concept, a complex deterrence posture of the sort needed to meet a wide variety of challenges will be unlikely to succeed; deterrence is expensive and time-consuming and requires a grand rationale if the necessary commitments are to be made and sustained.

That capstone concept is Integrated Deterrence. Integrated Deterrence is the optimal response to hybridised, cross-spectrum strategic challenges. It has four elements: vertical integration, horizontal integration, functional integration and temporal integration.

**Vertical Integration** concerns the coherence of all military components of deterrence, from the nuclear to the conventional and from the strategic to the operational and tactical levels of war, including the capability and effect of individual commanders and troops.

In brief, the argument for the maintenance of nuclear deterrence ideas, structures and capabilities can be made on two levels: ‘deterrence in principle’ and ‘deterrence in practice’. The first is to argue for nuclear deterrence as a valid and necessary concept, and for it to be maintained by established nuclear weapon states as a form of international stewardship. The possibility of further nuclear proliferation suggests that the long-standing taboo against nuclear weapons might be eroding, that the repugnance surrounding the prospect of nuclear use might not be shared by all states, and that nuclear use might therefore once again become thinkable. If such conceptual drift is indeed taking place, then it must be countered by the argument that nuclear weapons are – and can only be – a ‘non-weapon’ or ‘political weapon’. Absurd or self-evident as these two expressions might respectively appear, they lie at the heart of the argument for mutual nuclear deterrence. The second, ‘deterrence in practice’ is the relational argument, concerned with those circumstances in which only a nuclear deterrent capability could modify the assumptions and perhaps even the adversarial behaviour of another nuclear-armed power.22

But if a nuclear deterrent can be said to be a necessary component of an overall strategic deterrence posture, it cannot be said to be sufficient. As during the Cold War, if a strategic deterrence posture consists only of nuclear weapons then it is unlikely that anything below all-out nuclear war could successfully be deterred. Equally, without there being non-nuclear deterrence options, it is unlikely that nuclear deterrence would be considered credible in all circumstances.

A vertically integrated deterrence posture follows a circular logic which requires the posture to be capable, credible and communicable on all levels, not just the strategic nuclear. The rationale for nuclear deterrence is that it is part of a graduated response, a last resort once conventional deterrence has failed. And the rationale for conventional deterrence is that an adversary will understand that its bluff must not be called, precisely because there is a nuclear last resort. Yet NATO’s deterrence posture – on all levels – has been allowed to atrophy and as a result the vital connection between nuclear and conventional has been broken. The same cannot be said of Russia. With the recent articulation of the so-called ‘Gerasimov Doctrine’,23 the Russians have clearly thought very deeply about the vertical integration of their force posture, including the position of the nuclear threshold. They have modernized their command and control systems and their conventional warfighting capabilities. And they have also modernized their nuclear weapons, as evidenced by the provocative

---

22 This passage is taken from Paul Cornish, Perspectives on the Trident Successor Decision: Context, Purpose, Ethics and Cost, Submission to the Labour Party International Policy Commission, 8 June 2016: http://www.policyforum.labour.org.uk/.

deployment of SSC-8 cruise missiles, in contravention of the 1987 Intermediate-Range Nuclear Forces Treaty. NATO’s deterrence posture is left dangerously exposed by this turn of events.

NATO should urgently reconnect its nuclear and conventional deterrence capabilities, both doctrinally and in terms of capabilities. NATO’s 2010 Strategic Concept and its 2012 Deterrence and Defence Posture Review are both in need of refreshment. The Alliance’s nuclear weapon holdings and delivery systems should be subjected to an uncompromising re-evaluation. And on the conventional side, since NATO Europe no longer has the extensive military capabilities it once had, and given that European governments are persistently reluctant to increase levels of defence spending, ways must be found to use limited forces at the tactical and operational levels more selectively and cleverly, to bolster the strategic deterrent effect. For example, Anti-access/Area Denial (A2AD), described earlier as something to be prepared against (or ‘offset’) could be considered the basis of operational-level deterrence; more of an opportunity than a threat, in other words.

At the tactical level, more thought could be given to the idea of ‘devolved’ or ‘mobile’ deterrence as the modern equivalent of the ‘fleet in being’. The operating concept behind the ‘fleet in being’ was that a naval force, simply through knowledge of its existence, could exert a containing influence upon its adversaries, even without leaving port. In a similar way, the ‘army in being’ (or perhaps what some have called ‘Continuous On Land Deterrence’ – a deliberate emulation of the long-established and more plainly understood submarine-based deterrent posture known as ‘Continuous At Sea Deterrence’ (CASD)) would be required to demonstrate its existence, its capability and its determination without having to deploy the entire force. This might involve the very rapid deployment of troops and equipment, even at relatively low levels (such as a battlegroup), wherever and whenever required. But these deployments would not be in the sense of a human trip-wire whose task is simply to wait passively until being broken by the heavy footfall of an invading force. Instead, an ‘army in being’ would also involve the devolution of authority, decision-making and risk-taking down to the lowest level. Perhaps better understood as ‘deterrence by mission command’, the goal would be for the commander on the ground – the individual with the clearest sense of dangers and possibilities – to decide the most appropriate combination of military and non-military responses to a given challenge and in that way to connect the deterrent posture as closely as possible to the adversary’s actions and intentions.

**Horizontal Integration** contributes to the integrity of deterrence by ensuring that all relevant governmental and non-governmental bodies are included, as required, in the deterrent effort. Most, if not all NATO governments should be familiar with the claim that ‘joined-up’ government can improve policy effectiveness and efficiency and that public and private sectors should interact when required. In the UK these ideas are currently expressed in the form of the ‘Full Spectrum Approach’, as explained by Prime Minister David Cameron in his foreword to the 2015 National Security Strategy: ‘we will continue to harness all the tools of national power available to us, coordinated through the National Security Council, to deliver a “full spectrum approach”’. In March 2016 the goals of the Full Spectrum Approach were set out more fully in an informal briefing note:

---

*... to meet our priorities of deterring state-based threats, tackling terrorism, remaining a world leader in cyber security and ensuring we have the capability to respond rapidly to crises as they emerge. This push for flexibility and agility and the co-ordination of soft power, diplomatic, financial and other tools sits against the proliferation of new types of threats, including hybrid warfare and the rise of powerful non-state actors.*

---


26 NSS & SDSR 2015, p.6.

27 Unpublished informal briefing material.
On this basis it is easy to understand why some in and around UK government describe the Full Spectrum Approach as the UK’s ‘answer’ to hybrid warfare.

**Functional Integration** takes the broad principles of horizontal integration but applies them to functions and activities rather than to departments of government and other, non-governmental organisations. Activities, like organisations, must be deconflicted; trade sanctions, for example, should not be allowed unwittingly to undermine diplomatic negotiations. But functional integration should also seek to ensure that deconflicted activities can then be combined to *multiply* the overall deterrent effect. Nowhere is this more essential than in the field of information and communications technology (ICT).

In the UK, the ‘critical national infrastructure’ (CNI) is a subset of thirteen national infrastructure sectors. The CNI comprises specific assets and facilities, ‘the loss or compromise of which would have a major detrimental impact on the availability or integrity of essential services, leading to severe economic or social consequences or to loss of life.’ The CNI can be both physical (e.g. sites, installations, items of equipment) and logical (e.g. information networks, communications systems). A similar view informs the European Union Network and Information Security Directive:

> Network and information systems and services play a vital role in society. Their reliability and security are essential to economic and societal activities, and in particular to the functioning of the internal market.

Cyberspace – an unsatisfactory shorthand term for the medium of digital communication and information exchange – is increasingly vital to human activity on every conceivable level; political, economic, cultural and individual. It is also increasingly vital to strategic affairs. And what most distinguishes the 21st century strategic environment from its Cold War predecessor is without doubt the development of cyberspace as an arena not only for communication but also for strategic competition. We have thus come to the point where one of the essential ‘3Cs’ of deterrence – communication – has itself become a battleground. What is essential, therefore, is to ensure that cyberspace retains its functional integrity

Cyberspace can be protected offensively, through the ‘fusion’ of national levers of power in a form of punitive (and perhaps even anticipatory) deterrence. Strategic thinking seems to be shifting in this direction, evidenced by the more open use of such terms as ‘offensive cyber’ and ‘active cyber defence’. One of the most useful recent studies in this vein is that published by Kramer, Butler and Lotrionte:

> Offense will be a key element of effective operations. Prior to conflict, the United States should lead an expanded “fusion” effort, largely led by civil authorities, to bring to bear intelligence, cyber, financial, law enforcement, and other capabilities to disrupt the actions of state and state-associated entities undertaking adversarial cyber action.

In cyberspace, the contemporary equivalent of defensive deterrence (or deterrence by denial) is to ensure that ICT systems in and around the CNI remain resistant to intrusion and attack. Indeed, the Warsaw Summit Communiqué declared cyberspace to be ‘part of NATO’s core task of collective defence’ and recognised cyberspace ‘as a domain of operations in which NATO must defend itself as effectively as it does in the air, on land, and at sea.’ Very similar ‘deterrence by denial’ thinking can be found in the most recent UK National Cyber Security Strategy:

---

31 Warsaw Summit Communiqué, para. 70.
We will pursue a comprehensive national approach to cyber security and deterrence that will make the UK a harder target, reducing the benefits and raising the costs to an adversary – be they political, diplomatic, economic or strategic. We must ensure our capability and intent to respond are understood by potential adversaries in order to influence their decision-making. We shall have the tools and capabilities we need: to deny our adversaries easy opportunities to compromise our networks and systems; to understand their intent and capabilities; to defeat commodity malware threats at scale; and to respond and protect the nation in cyberspace.\(^{32}\)

As the foundation of Integrated Deterrence, cyberspace and ICT must be made resilient. The Concise Oxford Dictionary definition of ‘resilience’ is ‘recoiling; springing back; resuming its original shape after bending, stretching, compression etc.’ But in a world of fast-moving, asymmetric cyber threats, to ‘spring back’ to a position which has just been shown to be vulnerable to attack could, at best, be described as ‘dumb resilience’. Resilience must be more than resistance and restoration of the status quo ante. ‘Smart’ resilience is about ‘bouncing forward’ to a different, more advantageous position which allows the defender to regain some of the advantage just lost. This can be achieved, in part, through technical and managerial measures such as in-built redundancy and operational recovery plans.

But it might be possible for resilience to be ‘smarter’ still – perhaps even ‘dynamic’ – whereby the defender is more agile and adaptable and can regain and hold more of the initiative. Paradoxically, the winning move might be to think less about building a robust ‘adversary agnostic’ defence and more about identifying and managing vulnerabilities. It has been said that ‘in the information age, the people who know how to re-aggregate information will win.’ In pursuit of Integrated Deterrence this claim can be adapted as follows: ‘in the information age the people who know how to re-aggregate and redistribute vulnerability cannot lose.’ Dynamic resilience looks beyond the traditional model of deterrence as a largely static, stabilised relationship in which a limited set of known values (cities, ports, key communications nodes, military and logistic sites etc.) are held at mutual risk. Dynamic resilience allows the defender to reassess the value/vulnerability calculus to their advantage, whenever there is a case for doing so. In other words, the defender decides, on their terms, which risks they are prepared to take. And this assessment is made unilaterally; the idea of resilience is used simply as a framework to enable the defender to understand their own risk and vulnerability better than the adversary does.

The defender controls the resilience agenda by deciding on priorities and then retaining the ability and the initiative to change the analysis and reprioritise, as often as necessary. The starting point for dynamic resilience is for the defending side to have a clear and confident understanding of their own vulnerabilities and of the harms they might face in and from cyberspace.\(^{33}\)

Temporal Integration is an exercise in ‘future proofing’ Integrated Deterrence in order that it can be maintained over time and as circumstances change. National security and defence strategy is concerned with the passage of time. A national strategic outlook is rooted in a political, cultural and geostrategic past, perhaps several centuries old. The more recent past offers a bank of military, intelligence and other experiences, connecting to the deeper historical context and providing much needed ‘lessons’ in the application of national power – both successes and failures. National strategy is also concerned with the present, with articulating a view of the world, its strategic challenges and opportunities, and with organising national resources to achieve the optimal balance of capabilities – diplomatic, policing, development aid, military, trade, intelligence, cultural outreach and so on. And national strategy is also, of course, concerned with the future, particularly when it comes to the role of deterrence in shaping an aggressor’s cost-benefit calculations.

The obvious difficulty, for deterrence and for national strategy itself, is that the future is by definition unknowable and therefore unpredictable. Rafe Sagarin argued that adaptability is essential when preparing for a complex, uncertain and urgent future:


Adaptability is fundamentally different from merely reacting to a crisis (which happens too late) or attempting to predict the next crisis (which is almost certain to fail when complex ecological systems and human behaviors are involved). Adaptability controls the sweet spot between reaction and prediction, providing an inherent ability to respond efficiently to a wide range of potential challenges, not just those that are known or anticipated.34

One way to achieve adaptability in deterrence is to ensure that latency is written into strategic plans and preparations. Strategic latency is generally associated with technological innovation and development. But it is also a habit of mind whereby national strategy is forced to look beyond the short-term, to anticipate plausible strategic futures and to ensure that the long-range capability plan is not geared too tightly to a relatively short-term risk picture such as that contained within a government’s security and defence review. When financial scarcity bites, strategic resilience and adaptability can suffer; arguably the worst position to be in when threats and challenges seem to multiply at an unprecedented rate. Strategic latency is therefore an argument for national strategic resilience; for maintaining a balanced mix of armed forces and other levers of national power and influence even though (or, rather, because) such a posture might not appear consistent with a compressed and foreshortened assessment of national strategic priorities. Described in this way, latency is the foundation of a future-proofed, integrated deterrent posture.35


35 For further discussion of strategic latency see Cornish and Dorman, ‘Complex security and strategic latency: the UK Strategic Defence and Security Review 2015’.
Conclusion

Deterrence should continue to convey the message that the benefits expected from adventurism or aggression will be outweighed by the costs imposed. This expectation seems straightforward enough although, to paraphrase yet another Clausewitzian aphorism, ‘Everything in deterrence is very simple, but the simplest thing is difficult.’

The long-established principles and techniques of deterrence can (and must) be reconceived and modernised in ways which are more appropriate to 21st century European security challenges. This must be much more than an effort at strategic archaeology, however. Deterrence must not simply be rediscovered; it must be normalised with the European security discourse and on the broadest conceivable level, involving all functions and levers of government and including non-governmental bodies, the private sector and even individuals.

This paper has argued for Integrated Deterrence comprising four elements: vertical; horizontal; functional; and temporal integration. This is not intended to provide an exhaustive and exclusive account of 21st century deterrence. Far from it; Integrated Deterrence is best understood as a framework which can stand alone while allowing other ideas to be attached to it. If a broad understanding can be gained of the continued relevance and value of deterrence then the significance of other, perhaps more abstract ideas might also then be realised. There have recently been several noteworthy attempts to revive and reapply deterrence theory, such as extended deterrence;37 ‘triadic’ or indirect deterrence (i.e. deterrence of third party, non-state actors);38 normative deterrence (‘soft’ deterrence by the agreement on rules-based co-operation);39 associative deterrence (‘soft’ deterrence by the possibility of reputational damage);40 and ideational deterrence (or ‘deterrence by counter-narrative and delegitimisation’).

Some react with either lack of interest, suspicion or even open hostility to the idea that deterrence should be rehabilitated in any way whatsoever; surely it is a symbol of Europe’s uncomfortable, divided late 20th century past and as such is best understood as an historical artefact, and preferably not one to be excavated? Even senior diplomats and government officials appear on occasion to take something like this view. But in doing so they reveal both their lack of familiarity with their own history and their misunderstanding of deterrence as an idea. Deterrence is fundamentally not about the fighting of war in order to achieve the aim of national security. It is, instead, about the avoidance of war, in order to achieve that aim. Neither should this paper be interpreted as an attempt to ‘securitise’ (and least of all to ‘militarise’) European geopolitics, in the process returning Europe to its 20th century past. In the words of UK Foreign Secretary Boris Johnson, ‘Britain and its allies want no “new Cold War” with Russia.’ Instead, the aim of Integrated Deterrence is to answer asymmetric aggression and adventurism with a coherent, durable, symmetrical and adaptable strategic posture. Finally, Integrated Deterrence should not be perceived as an attempt to compensate for a strategic shortfall of some sort. Integrated Deterrence is not, in other words, a ‘Fourth Offset Strategy’ intended to mitigate the apparent, and alarming, lack of understanding of strategic matters within many Western governments. Integrated Deterrence is, instead, better understood as the ‘First Reset Strategy’, intended to refocus deterrence on its enduring purpose; the achievement of a stable and secure strategic environment.

36 Clausewitz, On War, p.119.
Recommendations

1. NATO’s strategic decision-making is too slow and tentative for the challenges the Alliance now faces. As well as having deterrent value in its own right, an adapted decision-making capability is essential if NATO’s strategic deterrence is to be both credible and communicable. NATO should therefore overhaul and re-energise its decision-making processes to be able to react to a fast-breaking crisis anywhere, at any time.

2. NATO should overhaul its command and control structures to enable the rapid deployment of credible forces anywhere around the Alliance’s periphery as well as beyond the NATO area as and when required.

3. NATO should revise and renew both the 2010 Strategic Concept and the 2012 Deterrence and Defence Posture Review, with the explicit goal of cementing the connection between nuclear and conventional deterrence capabilities.

4. Defence is a component of deterrence. In response to the Russian deployment of nuclear capable systems and weapons, NATO should pursue the closer integration of its air and missile defence systems. This initiative should seek to preserve the phased adaptive approach to the Middle East ballistic missile threat while ensuring that NATO develops the capability to defend against air and missile threats from Europe’s east.

5. NATO should reincorporate the c.180 US nuclear weapons (B61) deployed in Europe into a comprehensive and coherent nuclear policy. The B61 capability should undergo a modernisation programme to include storage and capacity. Nuclear Sharing countries should be required to confirm the readiness of their Dual-Capable Aircraft.