

NORAD and USNORTHCOM's Deterrence Conundrum
From Defence to Punishment to Denial
(not for citation or quotation without expressed permission of the authors)

Dr. Andrea Charron and Dr. James Fergusson
Centre for Defence and Security Studies
University of Manitoba
cdss@umanitoba.ca

The concept and theory of deterrence has long guided American/Western strategic thought and practice. Dating back to its origins following World War II, deterrence thinking and the complicated debates among academics and practitioners largely focused on one outcome and one assumption. First was the outcome of the impact of nuclear weapons married to ever more advanced delivery systems at the tactical, operational/theatre, and strategic level. Second was the context of the existential political struggle of the Cold War with the assumption that the United States was, and would remain, the leader of the western world as a status quo power.

While conventional, military considerations were not absent from strategic thought and practice, these considerations were intrinsically linked to the nuclear equation. Accepting Soviet conventional military superiority in the central front of Europe and the Soviet political objective to seize control of Western Europe and install pro-soviet communist governments was a given. American and NATO conventional forces served as a 'trip wire'¹ to communicate western resolve, alliance solidarity and to set in motion the process of escalation up the nuclear ladder to the American strategic nuclear deterrent. In other words, the idea that conventional forces alone could deter Soviet conventional aggression by denial was never relied upon. Rather, denial at whatever level possible, was simply a stepping-stone to deterrence by punishment.

Naturally, the place of conventional forces within the US-led western deterrence posture was hotly debated throughout the Cold War, especially in the 1960s around the American operational concept of Flexible Response.² Nonetheless, conventional forces or deterrence by denial (or what is now referenced as just defence) remained embedded under the umbrella of deterrence by punishment. There was always the fear that small events could trigger a nuclear exchange and so the quality of deterrence by denial was insignificant against the backdrop of mutually assured destruction.

¹ Some argue that a tripwire (bands of forward deployed conventional forces) is different from true deterrence by denial. In the case of the Cold War context, it wasn't just that small bands of forces could frustrate and slow the enemy down. The tripwire of the Cold War – hundreds of thousands of conventional forces forwardly deployed on both sides - was linked to the potential launch of nuclear weapons in retaliation. See Mitchell, A. Wess. 2015. "The Case for Deterrence by Denial", *The American Interest* (12 August). <https://www.the-american-interest.com/2015/08/12/the-case-for-deterrence-by-denial/>

² See Strometh, Jane E. 1988. *The Origins of Flexible Response: NATO's Debate Over Strategy in the 1960s*. London: Palgrave MacMillan. 1988.

Today, ideas about nuclear weapons have changed considerably. Indeed, one often forgets that a nuclear exchange is still a possibility. That China would articulate a no first-use nuclear strategy and others, including North Korea, adopt an “only in defence” policy speaks to the spread of a non-retaliatory norm that has stricken nuclear deterrence. For the United States, as the leading and most technologically advanced military power, deterrence by denial has become the dominant and preferred posture, ideally (or perhaps preferably) disconnected from nuclear weapons as a function of US conventional military superiority. But can one dismiss the punishment option from the calculus and how has the evolution of US deterrence doctrinal thinking applied to the North American context? This paper seeks to answer these last two questions by focusing specifically on the Canada-United States (CANUS) architecture and doctrinal thinking especially as it relates to the North American Aerospace Defense Command (NORAD). This analysis first sets the stage for evaluating NORAD and USNORTHCOM (N²) requirements by examining their place within broader political and strategic considerations related to the overarching US-led status quo deterrence posture. It then turns to identify and evaluate the factors related to N² requirements to evaluate the credibility of the North American deterrence by denial posture. First, however, it is useful to explore the shift from defence to deterrence facilitated by the creation of NORAD.

From Defence to Deterrence

Initially, Canada-United States (CANUS) homeland defence cooperation, which led to the operational establishment of NORAD in 1957, was largely informed by traditional defence thinking drawn from strategic bombing in World War II and the development of Soviet long range aviation (LRA) prior to the emergence of deterrence as the dominant concept and theory.³ Very quickly, however, NORAD’s primary mission shifted from defence *per se* to deterrence. Its denial function in this regard was to communicate to the Soviet Union that any attempt to *disarm* the United States through a pre-emptive first strike against its strategic command and control (C²) and/or nuclear retaliatory forces would fail. With the advent of intercontinental (ICBM) and submarine launched ballistic missiles (SLBMs), NORAD had no means to defend against such strikes. Rather, through its ballistic missile warning mission, NORAD provided sufficient warning through its integrated tactical warning/attack assessment (ITWA) process to provide the time for the United States National Command Authority (NCA) to order the release of its strategic nuclear retaliatory forces prior to Soviet first strike forces reaching their targets. This shift from defence to deterrence was evident with the significant reductions in fighter interceptors dedicated to air defence by both parties and the closure of ground-based surface-to-air (SAM) sites. Even with the development of air-launched cruise missiles in the 1970s, and with the aid of the new North Warning System (NWS) to provide advance notice, NORAD’s fundamental purpose in the deterrence equation was linked to the US punishment threat.

Today, NORAD, as well as USNORTHCOM and Canadian Joint Operations Command (N²+C), reside exclusively in the world of deterrence by denial, de-linked or arguably de-

³ See Goette, Richard. 2018. *Sovereignty and Command in Canada-US Continental Air Defence, 1940-1957*. Vancouver: UBC Press and Jockel, Joseph. 2007. *Canada in NORAD 1957-2007: A History*. Montreal and Kingston: McGill-Queen’s University Press.

coupled from deterrence by punishment as embodied in US strategic nuclear retaliatory forces. While the place of US strategic nuclear weapons or punishment forces within the North American deterrence by denial posture raises interesting questions in themselves, the key issues to consider are the role, function and requirements for a credible North American deterrent by denial posture, which in turn, would communicate to potential adversaries that they would be unable to achieve their objectives by threatening or attacking North America.

The Context of the North American Deterrent

Deterrence by denial is, by definition, a warfighting posture.⁴ It is designed to communicate to a potential adversary that any attempt to employ force to obtain its political objectives and thus revise, in some manner, the existing status quo, will fail. Traditionally, this meant a decision to use military force or war, and thus the objective of deterrence, in general and denial in particular, was to prevent war between the East and West. Today, however, this has expanded to cover a wide range of coercive tools short of war that may be employed to alter the status quo. Moreover, the status quo is no longer simply associated with territorial changes or territorial expansion *per se*, but also with political relationships and the existing world order, as created and maintained by the United States and its allies.

In this context, a credible deterrence by denial posture must be understood in multi-dimensional global terms. At its core, the North American deterrence by denial posture is fundamentally and existentially credible if North America can defend, or is perceived to be able to defend itself, if attacked. The questions become under what conditions and to what ends would an adversary decide to threaten explicitly or attack North America? Given that the People's Republic of China (PRC) and the Russian Federation (RF) are seen as the primary threats to the status quo today and for the foreseeable future as 'near peer' competitors, their objective would not be to provoke an existential conflict with all of its potential implications.⁵

Rather, the objectives of both states, as revisionist powers, are concentrated within a regional context, seeking to revise a regional balance of political power thus altering the world order from the bottom up. The United States has global status quo objectives and China and Russia regional revisionist objectives.⁶ In this sense, the United States has a global sphere of influence, whereas China and Russia have regional spheres of influence as core objectives, notwithstanding the future potential of either or both to act extra-

⁴ Rühle, Michael 2015. "Deterrence: what it can (and cannot) do". *NATO Review* (20 April) <https://www.nato.int/docu/review/articles/2015/04/20/deterrence-what-it-can-and-cannot-do/index.html>

⁵ At least in recent Canadian defence discussions, there are four plus one threatening actors – China, Russia, Iran, North Korea, plus non-state terrorism. This analysis concentrates exclusively on China and Russia, with a somewhat greater emphasis on Russia as a function of it being the primary concern for now of NORAD and USNORTHCOM notwithstanding the fact that China is the more likely aggressor from a world order perspective. (See Heather Conley et al. 2020). Iran, North Korea and non-state terrorism are in a distinct category from the 'near peer' competitors which this paper does not discuss.

⁶ See for example, the recent Rand Corporation study on deterrence and Russian 'redlines'. Pezard, Stephanie and Ashley Rhodes 2020. *What Provokes Putin's Russia? Deterring without unintended escalation*. Rand Corporation. January. <https://www.rand.org/pubs/perspectives/PE338.html>

regionally as a means to affect US leadership and credibility within their respective core regions. By revising or obtaining a dominant regional sphere of influence, both could alter the world order as it currently exists. In other words, both China and Russia seek to push the United States out of their respective regions not necessarily through the threat or use of force, but rather by actions short of direct confrontation in order to undermine American will and credibility in relation to its regional formal and informal allies.

The key point here is to understand that the N² deterrent cannot be seen and understood in isolation from the US-led global deterrent embedded in its forward regional command structure, deployed forces, and formal/informal allies. N² credibility is inherently linked to the integrity of the global deterrent. This, of course, has been the case at least throughout the era of the Cold War and the development of advanced nuclear weapons delivery systems. However, what has changed is the conceptual nature of this linkage.

During the Cold War, the conceptual emphasis, and thus priority for deterrence, was placed upon the forward regional commands. The North American deterrent component resided at the end of the global deterrent chain, evident in its place at the end of the vertical escalation ladder, and embedded in the need to ensure that US strategic forces would not become strategically or politically de-coupled from its forward components.

While the legacy of forward regional commands remains, the architecture logic has transformed requiring North America to be part of the calculus. The reasons for the change are five-fold. First, the challenges posed by America's (and by default Canada's) 'near peer' competitors are not existential in nature. Regional flashpoints or clashes of interests are not of the nature which would make a strategic retaliatory threat credible as was the case during the ideologically-driven Cold War. Second, conventional military superiority relative to potential regional flashpoints has arguably removed nuclear weapons from the global strategic deterrent posture. Whereas the purpose of nuclear weapons in the past was to deter a major war with existential implications, today their purpose is to deter the use of nuclear weapons.⁷ Third, new advanced long range delivery systems provide non-nuclear options for adversaries to threaten North America, notwithstanding the problem of knowing whether an advanced long range air-, sea- or ground-launched cruise missile possesses a conventional or nuclear warhead.⁸

Fourth, Russia and China have clearly demonstrated the will and ability to undertake coercive actions short of war that threaten the status quo, whether conceptualized

⁷ While strategic nuclear force modernization is on the US agenda, there is no indication that the current Administration will reverse President George H.W. Bush's 1991 decisions to withdraw forward deployed tactical and theatre nuclear forces (a series of decisions collectively known as the Presidential Nuclear Initiatives (PNIs)). Even with the collapse of the INF Treaty, it does not necessarily follow that either the US or Russia will proceed to re-introduce theatre nuclear forces.

⁸ During the final decade of the Cold War, it was simply assumed that the Soviet Union's first generation of air launched cruise missiles possessed nuclear warheads. Today, the US, at least publicly, differentiates types of cruise missiles on a conventional versus nuclear variant, of which, no doubt, the Russians and Chinese are fully aware.

regionally or globally. Alternatively labeled ‘grey zone’ conflict or ‘hybrid war’,⁹ these actions are calculated not to produce an American military response in the absence of an explicit threat from the United States. Finally, and arguably most importantly, globalization has generated a new set of vulnerabilities that did not exist in the past and can significantly affect the credibility of the US integrated global deterrent posture. These reside specifically in the economic and information (cyber) domains generating new and more complicated forms of homeland vulnerability that, in turn, affect the credibility of the US-led global deterrent as a status quo power.

In other words, North America’s place within the US global deterrent has been altered arguably to a priority heretofore not considered necessary. General O’Shaughnessy, the former commander of NORAD/USNORTHCOM (2018 - 2020), has regularly noted that the North American homeland is no longer a sanctuary.¹⁰ At one level, this is problematic as the homeland has never been a sanctuary *per se* once long range air and ballistic missile capabilities emerged. However, he is correct in the sense that it was largely thought of as a sanctuary in the past at the political and public levels, with the attendant implications that attention, investment and resources were better directed overseas to the regional commands.¹¹ Changing this thinking is the core deterrent requirement today, and with it altering investment and resource priorities. This, in turn, can be explained by deterrence theory.

It is important to recognize that deterrence, like war, is a two-way street. As Clausewitz pointed out: “So long as I have not overthrown my opponent I am bound to fear he may overthrow me. Thus, I am not in control: he dictates to me as much as I dictate to him”.¹² Similarly, in both dyads of concern, US-Russia and US-China, neither is in control *per se*, but each side is seeking to deter the other, and, in so doing, each dictates to the other. In other words, the US seeks to deter actions contrary to a regional status quo, while Russia and China respectively seek to alter this status quo by deterring the US from responding.

⁹ Conceptually, ‘hybrid war’ is best understood as a military subset of ‘grey zone’ conflict. In addition, ‘grey zone’ conflicts reside in political areas where vital American interests or clear signals of vital interests that could trigger a military response are absent. This, of course, has conceptual implications related to the applicability of deterrence. This occurred for example in the case of Crimea and Ukraine, which, at least, in terms of Morgan’s concept of immediate deterrence, indicates that Crimea and Ukraine were not a deterrence failure because the US did not practice deterrence. Arguably, the American deterrence dilemma resides in the realm of general deterrence. See Morgan, Patrick 1977. *Deterrence: A Conceptual Analysis*. Beverley Hills, CA.: Sage Publications. .

¹⁰ O’Shaughnessy, General Terrence J. 2019, “STATEMENT OF GENERAL TERRENCE J. O’SHAUGHNESSY, UNITED STATES AIR FORCE COMMANDER UNITED STATES NORTHERN COMMAND AND NORTH AMERICAN AEROSPACE DEFENSE COMMAND” Before the Senate Armed Services Committee (29 February) p. 2. https://www.armed-services.senate.gov/imo/media/doc/OShaughnessy_02-26-19.pdf

¹¹ Arguably, this was evident when the US decided roughly in the 1960s to forgo civil or passive defence measures, relying upon its offensive strategic retaliatory forces for defence. Germane to this point is the concept of Hardening the Shield (of North American defence). See O’Shaughnessy, Terrence and Scott Fesler. “Hardening the Shield: A Credible Deterrent and Capable Defense for North America”, Wilson Center (September 11, 2020). Found at <https://www.wilsoncenter.org/publication/hardening-shield-credible-deterrent-capable-defense-north-america>

¹² von Clausewitz, Carl. 1976. *On War*. Edited and translated by Michael Howard and Peter Paret. Princeton, New Jersey: Princeton University Press) p. 77.

In this context, the North American deterrence component consists of a threat of punishment by the adversary confronting the threat of denial. This punishment threat, in turn, can be conceptualized in two distinct ways in terms of adversarial calculations and potential actions at the regional level. First, if the benefit-cost calculation is perceived by the adversary in its favour, then it may be emboldened to undertake actions in its regional sphere of interest believing that this favourable calculation will deter the US from responding. In other words, a favorable calculation is a recipe for the adversary to act and generate a crisis condition believing that the US will not threaten escalation because of a vulnerability in the homeland. This is the existential condition that generates crises because no explicit threat is required by the adversary.¹³ The extant balance of offensive to defensive capabilities is sufficient to deter.

Second, in a crisis, this existential or implicit threat posture may then be translated into an explicit one in several ways. The adversary may explicitly communicate the nature of the threat and the conditions in which it will carry out the threat, publicly or privately. The adversary may undertake certain preparatory steps to signal a willingness to escalate. Finally, the adversary may undertake some actions outside of the military sphere, likely in the economic or information spheres, to signal a willingness to escalate further.

Whether implicit or explicit, US/allied decision-makers must calculate the value or importance of the specific issue, location, and nature of the challenge to the status quo (the benefits and costs of responding, threatening escalation or escalating) relative to homeland vulnerability.¹⁴ Of course, the adversary will make a similar calculation with regard to likely US/allied resolve. Regardless, given the dangers associated with crisis escalation, the preference of the US and its allies as the status quo actors is to deter challenges or crises from starting, rather than responding to the challenge. As such, the current North America balance of offensive (adversarial punishment capabilities, real or perceived) to defensive (North American denial capabilities, real or perceived) capabilities (or the deterrent difference) may generate incentives for adversaries to challenge the status quo at the local/regional level.

There are, of course, situations where crises occur not of the making of the adversary, but whose implications draw them in. While the tendency is to see every crisis as reflecting adversarial intent and an adversarial challenge, this is not always the case. Arguably, for example, Russia did not instigate the crises in Crimea, Ukraine, or Syria. It responded to them relative to its interests, which included a calculation of likely US responses. In such circumstances, the US did not practice immediate deterrence because the crises were

¹³ This is a variant of the current Russian strategic doctrine to ‘escalate in order to de-escalate’ in the context of a crisis.

¹⁴ Morgan, Forrest E., Karl P. Mueller, Evan S. Medeiros, Kevin L. Pollpeter and Roger Cliff. 2008. “The Nature of Escalation” in *Dangerous Thresholds: Managing Escalation in the 21st Century*. RAND Corporation. <https://www.jstor.org/stable/10.7249/mg614af.9> “Escalation can usefully be defined as an increase in the intensity or scope of conflict that crosses threshold(s) considered significant by one or more of the participants” p. 8

unexpected, the result of the actions of others and no explicit threat had been enunciated.¹⁵ Even so, this is an example of the US' general deterrence posture in a politically-peripheral area. It is unlikely that the North America component of general deterrence was significant in the events that followed, not least of all because it was Crimea and Ukraine in particular that fundamentally transformed the US-Russian relationship into an adversarial one. However, since the souring of relations, the North America component is likely to be significant in terms of future challenges (the failure of general deterrence) leading to future crises (immediate deterrence).

Overall, this is the new deterrent environment in which the North American component emerges as central or key. While both Russia and China are modernizing and developing new advanced military capabilities, both confront a superior global military power. Neither, at least for the time being, are likely to be willing to directly challenge US military superiority, but they will find ways to exploit inattention or complacency, especially in other domains including the information domain. Their interests are to challenge the status quo through actions that are perceived not to induce a US escalatory response. But, vulnerability in the homeland is something they can exploit to alter US calculations. The homeland is the weak link today in the US global general deterrence posture which will affect both the calculations and actions of Russia and China, and, perhaps most importantly, the calculations and decisions of the US (self deterrence).¹⁶ In other words, North America should be conceptualized as the 'hub' of its global deterrence posture, with the overseas regional commands as the spokes.

North America and N²+C Deterrence Requirements

Any evaluation of deterrence requirements must first recognize that the objective is North America, not Canada or the US separately *per se*. Their individual denial requirements cannot be separated. A threat to either is a threat to both. From this starting point, the current structure of the defence relationship underpinning a credible North American deterrence by denial posture is, itself, problematic. The relationship at its strategic and operational level is divided in several different ways with no overarching true central structure to provide unity of effort and command for North America. Part of the relationship is binational as embodied in NORAD with its functional responsibility for

¹⁵ In a somewhat similar case, the Iraqi decision to invade Kuwait was not an immediate deterrence failure as the US did not enunciate a specific threat to dissuade the Iraqi action.

¹⁶ This was the central argument of the Committee on the Present Danger in the 1970s in relation to the potential of the Soviet to undertake a successful strategic first strike as a function of its strategic force structure. The concern was not that the Soviets would actually undertake such a strike, but rather the US would back down in a crisis because of this perceived threat, thereby emboldening Soviet leadership and weakening the willingness of the US to respond. See Well, Samuel, Jr. 1979. "Sounding the Tocsin: NSC 68 and the Soviet Threat", *International Security*, Vol. 4, No. 2 (Fall), pp. 116-158.

aerospace (air and ballistic missile) and maritime warning¹⁷, and aerospace control (air). The remaining parts are bilateral.

Overall cooperation and coordination is located in the tri-command arrangement consisting of NORAD, USNORTHCOM and Canadian Joint Operations Command (CJOC) or N²+C established roughly a decade ago.¹⁸ It is, at best, an informal command arrangement, and whether it will evolve to become a more formal, centralized North American command remains to be seen. Moreover, the N² legs of the arrangement are devoted strictly to North America, whereas CJOC is responsible for all Canadian military operations, whether home or abroad, that do not involve NORAD, and continues to devote most of its attention and limited resources to overseas operations.¹⁹ (The recent COVID-19 epidemic and increasing requests for assistance to civilian authorities is changing this stance dramatically. Fifty percent of CAF operations are now conducted in Canada.) Nonetheless, to employ the language of the now defunct Binational Planning Group (BPG), N²+C, along with the mixed binational and bilateral components of the North American relationship, have resulted in creating North American command ‘seams’ with potential implications for deterrence credibility.

This, of course, is only one of many ‘seams’ of potential concern. Another seam is directly related to the concept of deterrence. This is the seam between denial and punishment. All of the North American command components strictly operate in the denial sphere. On the one hand, US punishment authority and capabilities relative to North America are assigned to another command within the US Unified Command Plan (UCP) – US Strategic Command (USTSRATCOM).²⁰ In contrast, the other regional commands in the UCP possess both denial and punishment authority and capabilities. On the other hand, Canada

¹⁷ In contrast to the exclusive aerospace warning mission, the maritime warning function is simultaneously binational and bilateral. For a detailed discussion of the complicated maritime warning relationship, see Charron, Andrea and James Fergusson. 2015. *Left of Bang: NORAD's Maritime Warning Mission and Maritime Domain Awareness*. Winnipeg: Centre for Defence and Security Studies. https://umanitoba.ca/centres/cdss/media/0_NORAD_Maritime_Warning_Mission_Final_Report_8_Oct_2015.pdf

¹⁸ To a lesser degree, the Permanent Joint Board on Defence (PJBD) and the Military Cooperation Committee (MCC) also provide input on cooperation requirements. The PJBD was established by the 1940 Ogdensburg Agreement (a one-page press release), and makes recommendations on defence cooperation to both national command authorities. The MCC in many ways is simply the technical arm of the PJBD. See Charron, Andrea. 2020. “The Permanent Joint Board on Defence; How Permanent and How Joint? Workshop Report (25 February). Winnipeg. Centre for Defence and Security Studies. Found at https://umanitoba.ca/centres/cdss/media/The-Permanent-Joint-Board-on-Defence-final-workshop-report_2020.pdf

¹⁹ For a brief period of time following the establishment of USNORTHCOM, Canada established a separate Canada Command, along with Canada Expeditionary Command, Canada Special Operations Forces Command (CSOFC), and Canada Canadian Operational Support Command (CANOSCOM). Primarily for cost reasons, and with the exception of CSOFC, these commands were merged into CJOC.

²⁰ USSTRATCOM, mirroring Russian behaviour, has resumed B-52 and B-1b bomber flights towards Russia without penetrating the Russian air defence identification zone. As a significant proportion of Russian GDP is generated from its eastern oil and gas fields, threatening them raises the costs for Russia relative to their threat to North America. Moreover, these fields are much more economically important to Russia than the North American Arctic is to either Canada or the US – notwithstanding the people who live there.

does not possess punishment capabilities. Certainly, such authority and capabilities could be given to NORAD, but successive Canadian governments have long ceded punishment to the US to be kept at arm's length from NORAD for political reasons, amongst others.

In terms of the US part of the deterrence equation, USNORTHCOM also confronts horizontal, geographical seams as a function of the UCP. It shares Alaska with US Indo-Pacific Command (USINDOPACOM). There are also three geographical seams in the other approaches to North America - the Atlantic and US European Command (USEUCOM), the Pacific and USINDOPACOM, and the south with US Southern Command (USSOCOM).

Alongside these seams, the North American deterrent also confronts domain seams. Reflecting, to some degree, military service structure, air, land, maritime and space remain conceptualized and structured into separate domains, even though it has been recognized that these domains increasingly blur together as a function of technological change. Thus, for example, a maritime threat as a function of cruise missile technology can quickly transition into an air breathing threat. In this regard, the US is currently investigating the concept of a Joint All Domain Command and Control (JADCC) structure. Its implications for the North American deterrent structure remains to be seen, but it does imply the potential merger of punishment and denial. Regardless of a long list of obstacles, ideally there needs to be some level of discussion and engagement with Canada in its development, rather than the traditional approach in which the US decides, and Canada is simply left to react.

There are two other domains that require consideration. One, not a traditional domain *per se*, is the terrorist world. Although terrorism has significantly dropped on the defence and security agenda, and questions arise as to whether terrorists can be deterred,²¹ this domain cannot be ignored as it resides in the seam between the military and civil security agencies. The other, which has risen significantly on the defence and security agenda is the cyber world and with it, the cognitive domain. In these worlds, denial and punishment are also separated (i.e. punishment in the cyber domain appears to be the exclusive domain of US Cyber Command) and entails both military, civilian security agencies and the private domain.

Beyond structural seams, there also exists significant capability deficiencies, or 'gaps' as identified in the 2005 BPG report. N² relative to its missions employs the concepts of detect, denial, defence and defeat. Although these are not necessarily conceptualized in a linear fashion, detect and defence are the key concepts to evaluate capability deficiencies.

²¹ See Wilner, Alex 2015. "Contemporary Deterrence Theory and Counterterrorism: A Bridge Too Far?" *International Law and Politics* Vol. 47: 439-462 and Davis, Paul and Brian Jenkins 2002. *Deterrence and Influence in Counterterrorism: A Component in the War on Al Qaeda*. RAND National Defense Research https://www.rand.org/content/dam/rand/pubs/monograph_reports/2005/MR1619.pdf. ; Morral, Andrew R. and Jackson, Brian A. 2009. *Understanding the Role of Deterrence in Counterterrorism Security*, RAND National Defense Research. https://www.rand.org/pubs/occasional_papers/OP281.html

In this regard, a credible capacity to detect and defend equates to a credible deterrence by denial posture.²²

Detection is the first ingredient of denial credibility and central to NORAD's mandate. NORAD's North American aerospace warning mission is essential as is its maritime warning mission but to a lesser degree given the complicated national and bilateral elements embedded in the process. In this regard, three key deficiencies stand out. First, NORAD's air warning component is almost exclusively defined as synonymous with the information provided by the North Warning System (NWS),²³ rather than from a North American Warning System (NAWS). Being as the NWS is technically obsolete, this represents a key failing. The NWS is unable to detect long-range air and sea launched cruise missiles. Although this deficiency is clearly recognized by all the relevant parties, and a binational structure is in place to identify sensor solutions as well as requirements to move and filter large amounts of sensor data into NORAD for analysis and action (NORAD modernization), there seems to be no pressing urgency to move forward.²⁴ Certainly, as the future system is likely to entail a complex of ground, air, maritime, and space-based sensors, technology hurdles do exist, especially in terms of systems' integration. The danger exists, however, that waiting for the final solution, rather than building the system as partial solutions come online will leave a major detection gap for sometime to come. Indicative, the current NWS radars reach the end of their life-cycle in 2025, and plans are underway to extend the radars until 2035 if not beyond.

Second, the future NWS/NAWS sensor system remains largely conceptualized as a perimeter system, looking outwards from the continent. While NORAD, in the wake of 9/11, acquired an internal air picture of North America through its link to Federal Aviation Administration (FAA) and NAVCANADA radars, it is unclear whether or not these internal radars possess a cruise missile detection and tracking capability and/or future improved drone-tracking technology. A perimeter system clearly needs to be augmented by internal detection capabilities, not least of all if the defence side of the equation was to fail at the perimeter.

Third, the detection domains remain largely separate, rather than integrated into an all domain detection and thus analysis structure. While NORAD does possess both an air and ballistic missile warning function, and with the latter, a space tracking function as well, these appear to be largely separate reflecting the traditional division between air and outer space. Yet, as the future NAWS will likely entail a significant space-based component, detection of threats against these key space-based assets is essential. Moreover, threats to

²² Defeat is conceptually problematic. On the one hand, it can be conceptualized as the activities in war, and thus after deterrence has failed. On the other, it can be conceptualized as simply synonymous with defence.

²³ The conceptualization is not just a Cold War 'hangover', but also a function of the longstanding agreement between Canada and the United States that NORAD's supporting architecture or infrastructure in Canada would be cost shared whereas US related infrastructure costs are solely borne by the US.

²⁴ Shaky economies as a result of COVID 19 will certainly not help move NWS modernization to the top of anyone's budget list.

these components also extend to a wide range of space-based assets vital to the military and the economy, especially in low earth orbit.²⁵

In other words, these threats are not just to these (especially space-based) assets, but are in fact threats to the North American homeland. Moreover, attacks against these assets are not just a physical attack against the territorial homeland, but could entail the direct loss of life. Adversarial calculations of such repercussions of attacks on assets only will be distinctly different from a direct threat or attack against North America. This does not imply that NORAD should acquire a space defence mission *per se* (see below for further discussion). Rather, that NORAD's ballistic missile warning mission should include threats against space-based assets as well as part of its integrated tactical warning/attack assessment function (ITW/AA).

In addition, the development of hypersonic glide weapon technology foreshadows the merger of space and air into a true 'aerospace' domain. Like the maritime domain, the ballistic threat of maneuverable hypersonics may transition into a maneuverable air threat. In other words, the space-missile and air domains need to be integrated into a single detection domain, along with the maritime domain to generate an integrated all domain North American common operating picture. As far as the final traditional domain, i.e. land, it is less pressing for inclusion not least of all for reasons of geography and three oceans that mitigate against a land invasion scenario.²⁶ The cyber domain, however, is vital to include.²⁷

Threats emanating from the cyber world have attracted significant and growing attention over the last several decades, and, for some time, the Air Forces (and to a lesser degree NORAD) have made a claim on the domain, notwithstanding US Cyber Command and its unclear role in the North American deterrence equation. Regardless, central to the detection problem in the cyber domain and distinct from the other domains, attribution of a cyber attack is extremely problematic. Due to the complexities of the internet, and the ability of states, such as China and Russia, to employ implicitly or explicitly private actors, it is difficult to ascertain whether any attack has been motivated just for the fun of it, for criminal purposes, and/or for state purposes. Moreover, this domain is structurally more complicated than the maritime domain, as it involves not only the military relative to its own systems and other government agencies, but also private actors within the economic system, especially with the overwhelming majority of critical infrastructure in private hands within the integrated North American economy. In this regard, private business interests related to the health of the company acts, to some degree, as a disincentive to even report cyber attacks.

²⁵ It is likely that this does not apply to the geo orbits. Being home to the US ballistic missile early warning assets and core military communications, any attack against geo orbit is likely to be interpreted as the first step to a strategic nuclear attack, and thus require a potentially pre-emptive response.

²⁶ That being said, civil unrest inside North America is a growing concern and cannot be ignored. Civil unrest, however, is likely to be managed under national command authority exclusively given clear constitutional limits in Canada and the U.S.

²⁷ The Mexican border issue is largely, if not exclusively, an immigration issue of little threat to North America, and is tracked by Canada for evidence of comparison to the US/Canada border.

Regardless, as long as North American officials continue to emphasize cyber vulnerabilities and fear the consequences, incentives exist for adversaries to exploit the cyber world, whether in real or imagined terms. In this regard, whether the attempt by Russia, as attributed, to influence the 2016 Presidential election had any real impact on its outcome, is a mute question. It is the attempt itself, and the fears it generated of other, potentially more devastating attacks, that is important to the Russian deterrent posture. At the core of this problem is detection and attribution.

A cyber attack occurs in near real time, with usually no warning. In contrast, the kinetic world provides, in varying degrees, early warning signals as a function of advanced intelligence and surveillance capabilities. One can expect, for example, that longstanding normal patterns of military activity will be significantly altered in preparation for employment (i.e. mobilization). This does not necessarily mean that a decision to use force has been made. In some cases, preparations may simply be a means of threat signaling to alter adversarial responses, with no intent to escalate to the use of force. There are also political contexts, which suddenly change or evolve over time that provide signals. Regardless, in the kinetic world, the probability or fear of a ‘bolt from the blue’ is greatly overstated.

Cyber is, however, a world of ‘bolts from the shadows’. As an element of deterrence, in this case by punishment, state-sponsored or directed deterrence attacks may simply be interpreted as a signal to demonstrate what an adversary can and might do in the future as a means to alter calculations. In other cases, it is simply to disrupt a state’s ability to track and react at a later point in the decision-making line or to obfuscate an adversary actions. Operating at a low level of effect, and thus having only a limited, temporally short marginal impact, such as shutting down a website, the act is meant to indicate the potential to do more damage. Moreover, at least to date, these attacks are calculated as insufficient to generate a kinetic response. Even more, as noted above, the problem of attribution and thus plausible deniability also adds complexity to the detection side of the equation. This is compounded further with the potential for embedded computer ‘viruses’, such as the case of stuxnet in Iran, that can be triggered under certain conditions, that may be undetected – a potential ‘attack in the making’.

Of course, political warning signals, in the case of an emerging challenge to the ‘status quo’, can be generated and transmitted across the complicated North American cyber world to require/promote greater vigilance. In addition, ongoing analysis to discern potential patterns of cyber attacks over time and space may provide some modicum of prediction, and thus detection. But, at the end of the day, detection is exclusively in the hands of the owners, private, public, and military, of each computer network. As a result, detection capabilities vary widely across the North American cyber world, and thus vulnerabilities also vary widely.

While one cannot expect every network in the North American cyber world to implement a common standard *per se*, it is vitally important that at least North American critical infrastructure adopt a common detection standard in terms of detection software, notwithstanding the problem of defining what is and is not critical infrastructure. In

addition, it is also essential that intelligence or information sharing be formalized across the private, public and military divides relative to cyber attacks after they have occurred.

In this regard, arguably, the state of the cyber domain in North America is reminiscent of the state of the intelligence world pre-9/11, and the maritime domain prior to the undertaking of significant steps in the following years, which would include NORAD's acquisition of its maritime warning mission, the creation of the National Maritime Integration Intelligence Office (NMIO) and Canada's Marine Security Operation Centres (MSOCs). In this regard, a NORAD, or perhaps N² cyber detection mission for North America might be conceptualized on the basis of maritime warning and its protocols. (See Figure 2). Designed not to duplicate existing and evolving private/public actors and processes, it would provide a centralized analytical function, based upon its ITW/AA, as the only North American 'eyes' at the end of the intelligence collection process as it currently exists nationally and bilaterally. In so doing, as NORAD has done with regard to maritime warning, it may also act to spur greater intelligence cooperation and information sharing across North American as a whole.

Alongside detection, defence is the second capability component of a credible North American denial deterrent. Like detection, gaps exist that may affect adversarial and North American (Canada and the US) deterrent calculations. Several stand out in the traditional defence domains. Assuming Canada agrees on a CF-18 replacement and given the US anti-cruise missile interceptors, the question becomes whether or not the intercept density relative to the assets assigned to NORAD are sufficient to deter cruise missile threats via defence? NORAD is also looking at existing northern forward operating locations (FOLs), as well as other possible locations farther south to meet maritime threats and potentially provide some form of layered defence. In addition, there is also a recognized requirement for in-flight refueling capabilities, and the need to consider the deployment of anti-cruise missile point defences.²⁸ All of these suggest strongly that more resources need to be dedicated to the air defence component of the North American deterrent, and, of course, with them, their integration into the detection side of the equation.

Related to air defence requirements, the aforementioned merger of air and space into a true aerospace domain raises issues of merging air and missile defence. This process is already underway as the US Army is developing an integrated air and missile defence battle command system. Of course, this raises the thorny issue of Canadian participation, and with it, issues concerning intercept priorities and centralized command and control, which derailed, in part, Canadian participation in ballistic missile defence (BMD) in 2004.²⁹ A reversal of the Canadian 'not yes' to missile defence could entail assigning command and control to NORAD. There is the remote possibility that this would trigger a merger of the J-3 position in the NORAD-USNORTHCOM command centre – the only position

²⁸ Reflecting this, in NORAD's Vigilant Shield 17 exercise, air defence personnel from the South Carolina National Guard were deployed to North Bay, Ontario.

²⁹ The formal Canadian decision not to participate would occur a year later. See James Fergusson. *Canada and Ballistic Missile Defence 1954-2009: Déjà vu all over again*. Vancouver: University of British Columbia Press. 2010.

currently not combined.³⁰ Canada's continued "not yes" to BMD is undermining the credibility of the North American denial posture; Canadian vulnerability provides a venue for an adversary to exploit.³¹

There are, of course, a range of other issues in this context. Assuming that the US proceeds with a third continental missile defence site in the US Northeast, its requirements may entail an advanced tracking and cueing radar deployed in Canada.³² Such a radar, in turn, would also likely serve other valuable detection functions related to North American defence.

Turning to the maritime domain, beyond the logic of evolving the current bilateral structure of the CANUS naval relationship into a binational one, the defence equation is problematic as a function of naval preferences towards forward defence against the cruise missile capable surface and sub-surface ships (the archers), rather than homeland defences against sea-launched cruise missiles (the arrows). While not ignoring the defence value of this preference, the archers are located outside both Royal Canadian Navy (RCN) and USNORTHCOM's AOR *per se*. Defence against the arrows is secondary, when it should be primary for North American deterrence. In this regard, major surface combatants (including the future RCN combat vessel) need to deploy sufficient anti-cruise missile air defences, and these defences need to be integrated into NORAD's air defence assets. At a minimum, the role of maritime assets need to be fully integrated into NORAD exercises to bolster North American deterrence requirements.

In the terrorist and cyber domains, defence has long been outside of the military mandate, and assigned the role of second responder to deal with the consequences of an attack, if necessary. Defence is in the hands of police forces, and bilateral cooperation between Canada and the US.³³ There appears no reason to change the military role *per se*, except to ensure that protocols governing the provision of mutual support are fully developed in response to a major incident. In this regard, the concept that has recently emerged is deterrence by resilience. Simply, this is the development of capabilities to mitigate the consequences of a major terrorist or cyber event quickly and effectively, thereby reinforcing deterrence credibility.

In many ways, deterrence by resilience is not a denial posture. Rather, it is a recognition that denial is not possible. In traditional military jargon, it is simply a damage limitation

³⁰ It is noted that the J3 USNORTHCOM position has many more missions than simply BMD and so it is more likely that the J3 positions will remain separate reflective of two different commands with different mission sets.

³¹ The extent to which the US would employ its current ground-based mid-course phase system, located in Alaska, and potentially a third site in the northeastern US, to defend Canada remains an open question, and is fraught with political and moral implications.

³² In the wake of the Canadian "no" in 2005, Raytheon officials privately queried whether the no would also extend to a radar site deployed at Goose Bay, Labrador. No answer was provided. James Fergusson. *Canada and Ballistic Missile Defence 1954-2009: Déjà vu all over again*. Vancouver: University of British Columbia Press. 2010.

³³ This is evident in the existence of International Border Enforcement Teams (IBETs) consisting of Canadian and American federal police who work together on both sides of the border.

posture, which, of course, does enhance credibility in demonstrating to an adversary that its attack will unlikely reap perceived benefits. It is in this regard, that both Canada and the US need to enhance their ability to provide assistance to civil agencies, and this assistance should not be constrained by the border, and, at a minimum, such requirements should be a priority for the tri-command structure.

Conclusion

From the perspective of North American homeland defence and security, the current CANUS command structure and capabilities are locked into an exclusive deterrence by denial posture. Punishment as an alternative is not an option, which does not mean that an adversary does not confront a credible punishment threat. Rather, the punishment threat, and thus punishment capabilities reside elsewhere, and are exclusively American. The question then is whether the CANUS part of the equation is adequately structured and resourced to present a credible denial threat to an adversary. Arguably, an adversary could be dissuaded from directly threatening or attacking independent of a punishment threat conceived as a last resort.

Importantly, any adversary, regardless of perceptions of denial credibility, cannot ignore or simply discount punishment, given the reality of US strategic conventional and military capabilities. Of course as a psychological theory designed to alter adversarial thinking and calculations, it is extremely difficult to know or predict how an adversary thinks and responds to a deterrence posture. Perhaps, then, what is more significant is how North American decision-makers think about their own credibility. It is here that the North American conundrum resides.

The North American component of the US-led western global deterrent posture should reside as the central deterrent hub, such that an adversary does not perceive it as a vulnerability that could be exploited to deter US-led responses to regional challenges. Yet, it is questionable whether or not US and Canadian decision-makers even think in these terms about its homeland. Both, arguably, remain fixated on the overseas components, with North America as an afterthought, despite the rhetoric.³⁴ Even more, beyond NORAD and USNORTHCOM, and to a much lesser degree CJOC, two different viewpoints exist – an American view that neither Russia nor China would dare strike at North America, not least of all because of its overarching superiority, and last resort strategic punishment capabilities, and a Canadian view which doesn't really think in deterrence terms, not least of all because it lacks the capabilities to deter credibly, and deterrence is an American responsibility.

³⁴ This is clearly reflected in the most recent defence report from the Congressional Research Service which devotes no direct attention to the homeland. Congressional Research Service. *Renewed Great Power Competition: Implications for Defence – Issues for Congress*. Washington: Congressional Research Service. April 7, 2020.

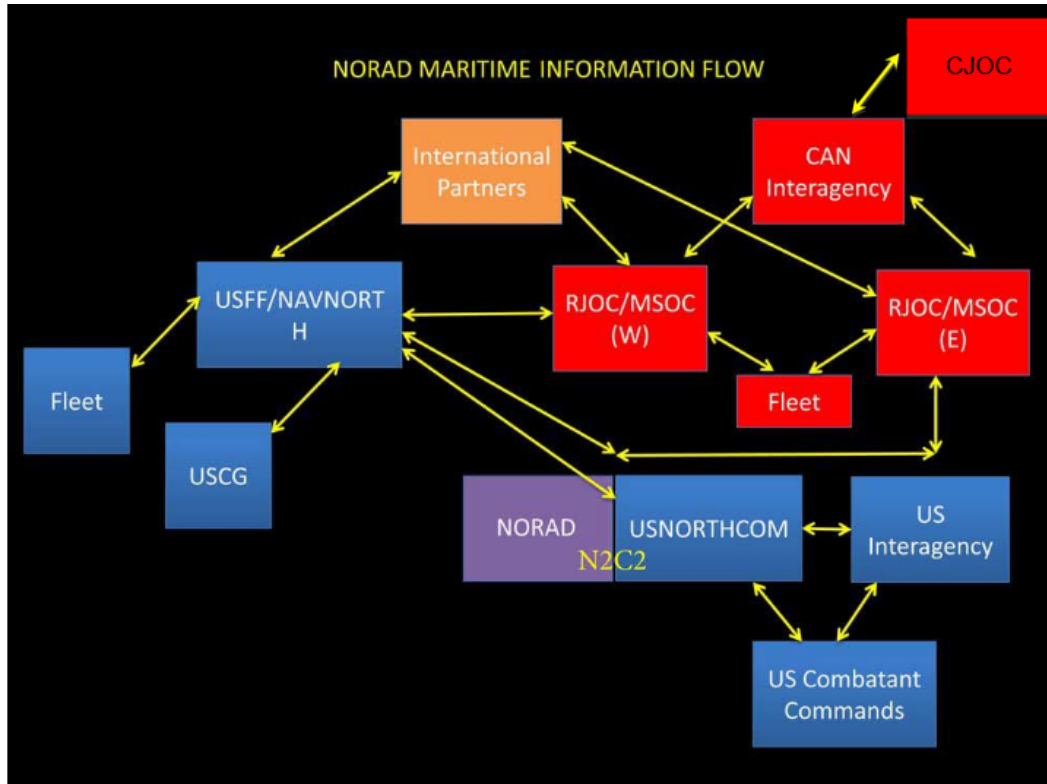
https://crsreports.congress.gov/product/pdf/R/R43838?utm_source=CSDS+News+and+Analysis&utm_campaign=23c3417fb9-EMAIL_CAMPAIGN_2020_04_16_07_30&utm_medium=email&utm_term=0_07827d2992-23c3417fb9-214634877&mc_cid=23c3417fb9&mc_eid=e3e01d6680

The net result may be a (vicious) feedback loop. An adversary comes to believe it can exploit homeland vulnerability, thus emboldening it to undertake a regional challenge by threatening actions short of war to deter a regional, overseas response. The US, and to a lesser degree Canada, quickly recognize their vulnerability (as well as vital overseas allies), and are unwilling to respond effectively, being forced to fall back on a strategic punishment threat to deter. This, in turn, further emboldens the adversary to undertake further challenges which raises doubts among overseas allies that the US will defend them.

The basic answer is to alter deterrence thinking in North America. It is vital that the structural changes to the North American deterrence posture and the necessary investments to alter adversarial perceptions are made so that North America cannot be held hostage. In other words, the current North American deterrence by denial posture remains embedded in an outdated Cold War mindset that has largely evolved in an *ad hoc* manner. Of course, this is easier said than done, and despite best efforts by senior NORAD and USNORTHCOM officials to communicate this message, it may take an unexpected overseas regional challenge that produces a major crisis in which the lack of North American denial credibility comes to the fore. Unfortunately, by then, it may be too late. The need to refocus on denial is paramount.

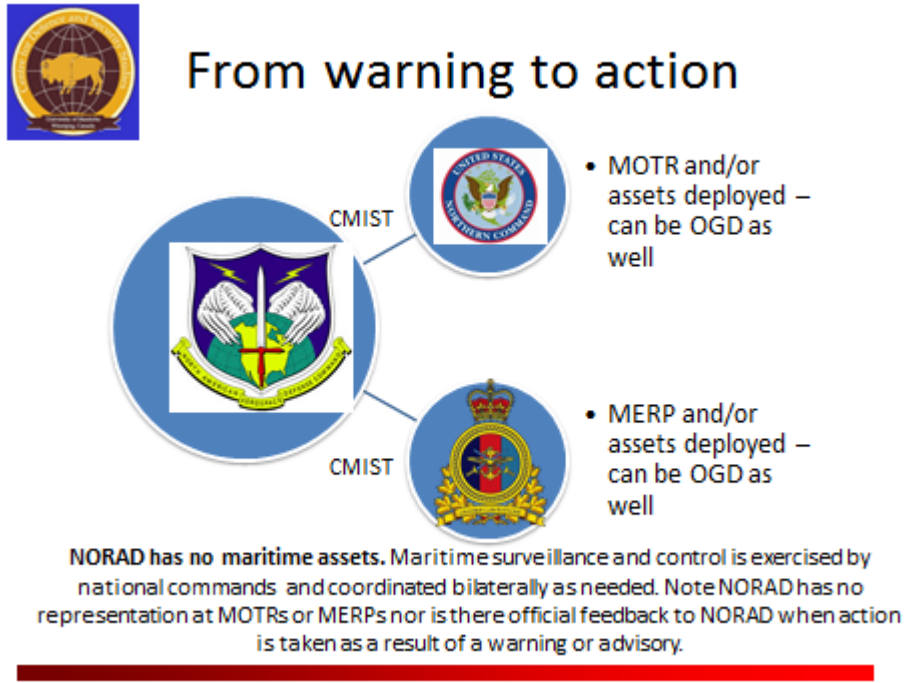
Annex

Figure 1- NORAD Maritime Information Flow³⁵



³⁵ Figure 1 shows a simplified model of the formal links that allow intelligence and information to be gathered to create the binational North American COP that is assembled at USFFC/NAVNorth and fed into USNORTHCOM to NORAD via N2C2. Each arrow seen in Figure 1 has its own filter that will add/delete information and the arrows can go in both directions. These filters include numerous factors which relate to organizational interests, mandates, jurisdictions and responsibilities and broader national interests. Even on the defence-side of the equation, NORAD's picture is a filtered one, especially relative to U.S. strategic interests and requirements, which differ significantly from Canada, and may be outside the purview of USNORTHCOM as well

Figure 2: Maritime Warning initiated at NORAD³⁶



³⁶ CANUS Maritime Information Sharing Teleconference (CMIST): is a telephone notification system that allows NORAD to contact relevant agencies on both sides of the border, which can vary depending upon the threat scenario. The CMIST process is not directly linked into the two national response structures: the Canadian Maritime Emergency Response Protocol (MERP), and the U.S. Maritime Operational Threat Response (MOTR).

Bibliography

- Charron, Andrea and James Fergusson. 2015. "Left of Bang: NORAD's Maritime Warning Mission and North American Domain Awareness" (8 October) Winnipeg: Centre for Defence and Security Studies.
https://umanitoba.ca/centres/cdss/media/0_NORAD_Maritime_Warning_Mission_Final_Report_8_Oct_2015.pdf
- Charron, Andrea. 2020. "The Permanent Joint Board on Defence; How Permanent and How Joint? Workshop Report (25 February). Winnipeg. Centre for Defence and Security Studies. Found at https://umanitoba.ca/centres/cdss/media/The-Permanent-Joint-Board-on-Defence-final-workshop-report_2020.pdf
- Congressional Research Service. *Renewed Great Power Competition: Implications for Defence – Issues for Congress*. Washington: Congressional Research Service. April 7, 2020.
https://crsreports.congress.gov/product/pdf/R/R43838?utm_source=CSDS+News+and+Analysis&utm_campaign=23c3417fb9-EMAIL_CAMPAIGN_2020_04_16_07_30&utm_medium=email&utm_term=0_07827d2992-23c3417fb9-214634877&mc_cid=23c3417fb9&mc_eid=e3e01d6680
- Conley, Heather, Nikos Tsafos and Ian Williams. 2020. "America's Arctic Moment: Great Power Competition in the Arctic to 2050" (30 March) Washington D.C. Center for Strategic and International Security. <https://www.csis.org/analysis/americas-arctic-moment-great-power-competition-arctic-2050>
- Davis, Paul and Brian Jenkins 2002. *Deterrence and Influence in Counterterrorism: A Component in the War on Al Qaeda*. RAND National Defense Research https://www.rand.org/content/dam/rand/pubs/monograph_reports/2005/MR1619.pdf.
- Fergusson, James. 2010. *Canada and Ballistic Missile Defence 1954-2009: Déjà vu all over again*. Vancouver: University of British Columbia Press.
- Goette, Richard. 2018. *Sovereignty and Command in Canada-US Continental Air Defence, 1940-1957*. Vancouver: UBC Press.
- Jockel, Joseph. 2007. *Canada in NORAD 1957-2007: A History*. Montreal and Kingston: McGill-Queen's University Press.
- Mitchell, A. Wess. 2015. "The Case for Deterrence by Denial", *The American Interest* (12 August). <https://www.the-american-interest.com/2015/08/12/the-case-for-deterrence-by-denial/>

Morgan, Forrest E., Karl P. Mueller, Evan S. Medeiros, Kevin L. Pollpeter and Roger Cliff. 2008. "The Nature of Escalation" in *Dangerous Thresholds: Managing Escalation in the 21st Century*. RAND Corporation. <https://www.jstor.org/stable/10.7249/mg614af.9>

Morgan, Patrick 1977. *Deterrence: A Conceptual Analysis*. Beverley Hills, CA.: Sage Publications.

Morrall, Andrew R. and Jackson, Brian A. 2009. *Understanding the Role of Deterrence in Counterterrorism Security*, RAND National Defense Research. https://www.rand.org/pubs/occasional_papers/OP281.html

O'Shaughnessy, General Terrence J. 2019, "STATEMENT OF GENERAL TERRENCE J. O'SHAUGHNESSY, UNITED STATES AIR FORCE COMMANDER UNITED STATES NORTHERN COMMAND AND NORTH AMERICAN AEROSPACE DEFENSE COMMAND" Before the Senate Armed Services Committee (29 February) . https://www.armed-services.senate.gov/imo/media/doc/OShaughnessy_02-26-19.pdf

O'Shaughnessy, Terrence and Scott Fesler. "Hardening the Shield:A Credible Deterrent and Capable Defense for North America", Wilson Center (September 11, 2020). Found at <https://www.wilsoncenter.org/publication/hardening-shield-credible-deterrent-capable-defense-north-america>.

Pezard, Stephanie and Ashley Rhodes 2020. *What Provokes Putin's Russia? Deterring without unintended escalation*. Rand Corporation. January. <https://www.rand.org/pubs/perspectives/PE338.html>

Rühle, Michael 2015. "Deterrence: what it can (and cannot) do". *NATO Review* (20 April) <https://www.nato.int/docu/review/articles/2015/04/20/deterrence-what-it-can-and-cannot-do/index.html>

Strometh, Jane E. 1988. *The Origins of Flexible Response: NATO's Debate Over Strategy in the 1960s*. London: Palgrave MacMillan. 1988.

von Clausewitz, Carl. 1976. *On War*. Edited and translated by Michael Howard and Peter Paret. Princeton, New Jersey: Princeton University Press.

Well, Samuel, Jr. 1979. "Sounding the Tocsin: NSC 68 and the Soviet Threat", *International Security*, Vol. 4, No. 2 (Fall), pp. 116-158.

Wilner, Alex 2015. "Contemporary Deterrence Theory and Counterterrorism: A Bridge Too Far?" *International Law and Politics* Vol. 47: 439-462.