
Dancing with the Bear

Managing Escalation in a Conflict with Russia

In collaboration with the Atomic Energy Commission (CEA)

Forrest E. Morgan

Winter 2012



Security Studies Center

The Institut Français des Relations Internationales (Ifri) is a research center and a forum for debate on major international political and economic issues. Headed by Thierry de Montbrial since its founding in 1979, Ifri is a non-governmental, non-profit organization.

As an independent think tank, Ifri sets its own research agenda, publishing its findings regularly for a global audience.

Using an interdisciplinary approach, Ifri brings together political and economic decision-makers, researchers and internationally renowned experts to animate its debate and research activities.

With offices in Paris and Brussels, Ifri stands out as one of the rare French think tanks to have positioned itself at the very heart of European debate.

The opinions expressed in this text are the responsibility of the author alone.

Thérèse Delpech (1948 – 2012)

Thérèse Delpech passed away on January 18, 2012. As Director of Strategic Affairs of the French Atomic Energy Commission (CEA), Thérèse was instrumental in promoting and supporting several research programs on proliferation in France and abroad. But for her and her continuous support along the years, the *Proliferation Papers* would not exist. Ifri's Security Studies team is as sad as we are indebted to her, and would like to dedicate the 2012 issues of the *Proliferation Papers* to her memory.

ISBN : 978-2-86592-999-3

© Ifri – 2012 – All rights reserved

Ifri
27 rue de la Procession
75740 Paris Cedex 15 – FRANCE
Tel : 33 (0)1 40 61 60 00
Fax : 33 (0)1 40 61 60 60
Email : ifri@ifri.org

Ifri-Bruxelles
Rue Marie-Thérèse, 21
1000 – Brussels – BELGIUM
Tel : 32 (0)2 238 51 10
Fax : 32 (0)2 238 51 15
Email : info.bruxelles@ifri.org

Website : <http://www.ifri.org/>

Winter 2012

***Dancing with the Bear:
Managing Escalation in a
Conflict with Russia***

Forrest E. Morgan

Proliferation Papers

Though it has long been a concern for security experts, proliferation has truly become an important political issue with the last decade, marked simultaneously by the nuclearization of South Asia, the weakening of international regimes and the discovery of frauds and traffics, the number and gravity of which have surprised observers and analysts alike (Iraq in 1991, Libya until 2004, North Korean and Iranian programs or the A. Q. Khan networks today).

To further the debate on complex issues that involve technical, regional, and strategic aspects, Ifri's Security Studies Center organizes each year, in collaboration with the Atomic Energy Commission (CEA), a series of closed seminars dealing with WMD proliferation, disarmament, and nonproliferation. Generally held in English, these seminars are structured around the presentation of an international expert.

Proliferation Papers is a collection, in the original version, of selected texts from these presentations. An anonymous peer-review procedure ensures the high academic quality of the contributions. Download notifications are sent to an audience of several hundred international subscribers upon publication.

Editorial board

Editor: Etienne de Durand

Deputy Director: Corentin Brustlein

Principal Scientific Adviser: Jean Klein

Layout Assistant: Alexandra Vickery

How to cite this publication

Forrest E. Morgan, "Dancing with the Bear. Managing Escalation in a Conflict with Russia", *Proliferation Papers*, No. 40, Winter 2012.

About the Author

Forrest E. Morgan is a senior political scientist at the RAND Corporation and an adjunct professor at the University of Pittsburgh Graduate School of Public and International Affairs. His research at RAND has addressed a variety of military strategy and doctrine issues regarding such topics as deterrence and escalation management, crisis stability, U.S. Army information operations, America's dependence on space, and assessing the performance of the U.S. Air Force and U.S. Army in Operation *Iraqi Freedom*. He is the author and co-author of several books, including: *Deterrence and First-Strike Stability in Space: A Preliminary Assessment* (RAND, 2010), *Dangerous Thresholds: Managing Escalation in the 21st Century* (RAND, 2008), and *Compellence and the Strategic Culture of Imperial Japan: Implications for Coercive Diplomacy in the 21st Century* (Praeger, 2003).

Before joining RAND in 2003, Dr. Morgan served a 27-year career in the U.S. Air Force. His assignments included duty as a signals intelligence analyst and as a space operations officer in various operations and staff positions. In the latter phase of his military career he served on the strategy and policy staff at Headquarters, U.S. Air Force, Pentagon, and as professor of comparative military studies at the Air University School of Advanced Air and Space Studies. He holds a PhD. in policy studies from the University of Maryland and a master of airpower arts and sciences degree from the Air University School of Advanced Air and Space Studies.

The views expressed in this paper are the author's alone and do not represent those of the RAND Corporation, the U.S. Air Force, or any U.S. government agency.

Contents

Introduction	7
Escalation Management during the Cold War	11
The Fundamental Question: Can Nuclear War Be Limited?	11
Herman Kahn and Escalation Dominance	13
Thomas Schelling and Brinkmanship	16
The Cold War Escalation Management Strategy of Choice: Conflict Avoidance	17
A New Approach to Escalation Management	19
A Closer Look at Escalation	19
The Nature of Escalation Thresholds	20
Understanding Mechanisms and Motives	22
The Importance of Understanding the Balance of Interests	28
Managing Escalation Risks In A Conflict With The Russian Federation	33
Potential Routes to Conflict between Russia and NATO	33
Escalation Dynamics in a Confrontation between NATO and Russia	35
Managing Escalation Risks in a Confrontation with Russia	40
Conclusion	45

Introduction

“Escalation”, the tendency of belligerents to increase the force or breadth of their attacks to gain advantage or avoid defeat, has characterized wars throughout history. Prussian military theorist Carl von Clausewitz had this dynamic in mind when he proposed that war, being a contest between interacting human beings would, in theory, culminate in each opponent’s maximum exertion of strength.¹ Although the dynamics of escalation were thus recognized as early as the beginning of the 19th century, a body of theory on how to manage it did not emerge until the Cold War, when the nuclear capabilities of the superpowers threatened to make the costs of uncontrolled escalation in any military conflict between them horrific. As nuclear arsenals grew in the 1950s, and especially after the 1962 Cuban missile crisis underscored the risks that a superpower confrontation might result in nuclear war, policy makers and security analysts put a great deal of thought into how to manage escalation in order to mitigate those risks. Such efforts should have persisted into the post-Cold War era – after all, nuclear weapons still existed, and tendencies to escalate limited conflicts remained a part of human nature – but, unfortunately, for a considerable period of time they did not.

For about 15 years after the Cold War ended, Western military and political leaders gave little thought to escalation risks in conflicts between states or with other dangerous actors in the global environment. There seemed to be no need. Although Washington and Moscow still had immense nuclear arsenals at their disposal, absent the ideological struggle of the bygone era, the risk that some new confrontation would take them to the brink of war seemed negligible, and no other state had nuclear capabilities that even approached those of the former Cold War rivals. In the conventional realm, the United States and its Western partners had demonstrated overwhelming military superiority in the first Gulf War and several subsequent operations, revealing that convergent advances in technology, organization, and doctrine had granted the West unparalleled superiority in joint, high-speed warfare. To U.S. leaders, it appeared that the West would have clear dominance in any future conflict at both conventional and nuclear levels, so escalation would not be a serious concern. Future adversaries would be cowed. Wars, if they did occur, would be fought at whatever level the Western powers chose.

Yet even early in the post-Cold War era there were indications that escalation management should remain an important consideration in all military operations. In 1991, as coalition forces postured to eject Iraqi forces

¹ Carl von Clausewitz, *On War*, Princeton, Princeton University Press, 1976, p. 77.

from Kuwait, concerns arose that Saddam Hussein might use chemical weapons in the imminent conflict. He did not, presumably, because he was deterred by U.S. Secretary of Defense Dick Cheney's implied threat to punish such an act with a nuclear response. Once the air campaign was underway, however, Saddam found another convenient escalation pathway through SCUD missile attacks against Israel, which threatened to broaden the conflict and, at the same time, fracture the coalition.

Other conflicts brought new surprises. In 1993 the United States discovered that its conventional military superiority did little good in Mogadishu, where warlords frustrated U.S. participation in a UN nation-building effort by escalating irregular warfare to a level at which the costs were unpalatable to the U.S. public. Later in the decade, Serbian combatants in Bosnia and Kosovo discovered they could escalate conflicts in ways that created new dilemmas for their opponents and the international community, with systematic rape, holding UN safe areas hostage, and ethnic cleansing. Finally, the terrorist attacks on September 11, 2001, and developments since have revealed that non-state adversaries can be remarkably resourceful in finding ways to escalate conflicts with powerful states.

By 2004, strategic planners at Headquarters U.S. Air Force had become concerned that they did not adequately understand escalation risks in the contemporary security environment.² Consequently, then-Air Force Chief of Staff General John P. Jumper tasked the RAND Corporation to examine the 21st century security environment for possible escalation risks and offer recommendations on how the Air Force could best manage any that might be found. In the study that followed, RAND determined not only that significant escalation risks did indeed exist, but that escalation management concepts developed during the Cold War would probably be inadequate for managing those risks. The study observed that, whereas Cold War escalation management approaches focused solely on managing confrontations between nuclear superpowers, new methods would be needed to manage risks in a security environment that had become much more complicated, with potential adversaries falling into three relatively distinct but interrelated categories: large nuclear powers, such as China and Russia; new and emerging regional nuclear powers, such as North Korea and Iran; and transnational networks of insurgents, terrorists, and criminals. The study then proceeded to examine the escalation dynamics

² Ironically, their concerns arose from observations other than those mentioned above. An increasing number of war games conducted by the various military staffs in the Pentagon since the late 1990s had ended in uncontrolled escalation, games in which the scenarios called for only limited U.S. military intervention against notional adversaries that were clearly outmatched by U.S. forces. How could such operations get out of control? At first game analysts assumed the outcomes were spurious, the result of overly aggressive "red teams", or perhaps the advanced systems being postulated in some of the futuristic scenarios were somehow escalatory by nature. But the increasing frequency with which the games turned escalatory and the wide range of participants and scenarios involved suggested something else was at work, something that Air Force planners did not understand.

that might arise in conflicts between the United States and opponents in each of those categories and offer recommendations on how to manage the risks that such conflicts would present.³

This paper builds on that work and makes a modest attempt to fill a gap left by it. When examining the escalation risks in conflicts with large nuclear powers, the RAND team focused exclusively on potential confrontations with the People's Republic of China. Reasons for this narrow focus can be traced mainly to client interest the year the study was conducted and the limited resources available to do the work. Yet there is another unspoken reason why the RAND study did not address the risks of escalation in a confrontation with Russia and, in fact, relatively little work has been done on Russia-related security issues in most Western research institutions since the end of the Cold War: Western analysts tend to undervalue the probability of future war between the Russian Federation and the West and, to the extent that they concede the possibility exists, discount Russia as a credible adversary.⁴

Such thinking – or perhaps lack of thinking – constitutes a blind spot in Western security analysis that is potentially dangerous. Granted, the intense rivalry that characterized the Cold War ended a generation ago, and Moscow is now on amicable terms with Washington and other North Atlantic Treaty Organization (NATO) capitals, but Russian interests could once again come into conflict with those of the West. Russia is a hobbled great power, struggling to regain its status in the geopolitical order and resentful that NATO has pushed into its traditional sphere of influence.⁵ While its conventional military capabilities are but a faint shadow of those Moscow wielded during the Soviet era – which is the reason Western analysts tend to dismiss Russia as a threat – it still maintains more nuclear weapons than any other country in the world. Should a future conflict of interests result in a military confrontation, the volatile dynamics that could result from this combination of conventional vulnerability and nuclear strength is reason for concern, not indifference.

This paper explores the challenges that NATO would face in managing escalation in a military conflict with the Russian Federation. It argues that, to manage escalation effectively, Western military and political

³ In 2008 RAND made the final report from this study publically available as a published monograph. See Forrest E. Morgan, Karl P. Mueller, Evan S. Medeiros, Kevin L. Pollpeter and Roger Cliff, *Dangerous Thresholds. Managing Escalation in the 21st Century*, Santa Monica, RAND Corporation, 2008.

⁴ As later references in this paper will indicate, RAND has actually done substantial research on Russia-related issues in the post-Cold War era, but that work has focused mainly on economic concerns and foreign policy questions.

⁵ Olga Olikier, Keith Crane, Lowell H. Schwartz and Catherine Yusupov, *Russian Foreign Policy. Sources and Implications*, Santa Monica, RAND Corporation, 2009, pp. 83-111; Michael J. Dean and Margaret A. Harlow, *Russia Workshop*, report from the Strategic Assessments Office, National Security Analysis Department, Johns Hopkins University Applied Physics Laboratory, Laurel, Johns Hopkins University, May 10, 2007, pp. 6-8, available at: http://www.jhuapl.edu/ourwork/nsa/workshop_files/Russia_GA.pdf, accessed: June 21, 2011.

leaders would need to understand the mechanisms and motives that drive it and the dynamics that could emerge when those mechanisms and motives engage during a confrontation. Equally important, however, they would need to weigh their interests in the issue at hand, vis-à-vis Moscow's, and adjust their war aims and efforts accordingly. Escalation management is about keeping limited wars limited. Consequently, it confronts the fundamental paradox of limited war: how to "win" while restraining one's efforts.⁶ The solution to that paradox lies largely in how one defines winning. In a war against Russia, Western leaders would want to seek victory. They could do so only to the extent that victory is defined and pursued in ways that ultimately allow for compromise and do not threaten the survival of the Russian state or its leaders.

The paper begins with a brief examination of the nature of escalation and a review of past approaches for managing it. Explaining the shortcomings of those concepts, it then proposes a different approach, one based on *threshold management*: that is, by more clearly illuminating each side's stakes and escalation thresholds and deliberately manipulating escalation mechanisms to keep conflicts safely within those thresholds. The next section explores the escalation risks that would arise in a conflict with Russia and describes how threshold management concepts might be applied to mitigate those risks. Finally, the paper compares the findings of this analysis with those of the RAND study, draws some additional insights about the strengths and limitations of threshold management, and proposes avenues for further research.

⁶ As Clausewitz argued in his unfinished magnum opus, *On War*, wars in the real world, as opposed to war in theory, are limited by a host of factors such as uncertainty, friction, constraints of time and space, defensive advantages, and, of course, the limited nature of political objectives (see Clausewitz, *On War*, *op. cit.*, pp. 78-86). In a July 1827 note regarding his plans to revise the draft text, he expressed his desire to clarify the point that wars can be of two kinds, either to overthrow the enemy, making him politically and militarily impotent, "*or merely to occupy some of his frontier-districts so that we can annex them or use them for bargaining at the peace negotiations*" [*Italics in original.*] (p. 69). Not until the advent of nuclear weapons, however, did theorists give serious thought to how to deliberately limit wars to reduce the risks of catastrophic destruction. As this essay will later show, a rich body of literature on limited war emerged in the latter half of the 1950s, largely in response to the Eisenhower administration's massive retaliation doctrine. This literature was overlapped and, to some degree, superseded by that on escalation management, which emerged in the 1960s. Like the latter, work on limited war declined at end of the Cold War. However, it has recently begun to revive with discussions on the changing nature of war and whether European armies are appropriately sized, organized, and equipped for contemporary conflicts. The Cold War-era limited war literature is briefly summarized later in this essay. For a couple of the more important recent publications on limited war, see: Hew Strachan and Sibylle Scheipers, *The Changing Nature of War*, Oxford, Oxford University Press, 2011; and Hew Strachan, "Are European Armed Forces Only Able to Wage Limited Wars?", Paris, French Institute of International Relations, Spring 2011, available at: <http://www.ifri.org/downloads/hewstrachanpe22011.pdf>.

Escalation Management during the Cold War

“Escalation” is a term used to describe a substantive increase in the intensity or scope of conflict. As described in the RAND study:

It is a fundamental dynamic in which adversaries engaged in a contest for limited objectives increase the force or breadth of their attacks to gain advantage or avoid defeat. Escalation can be unilateral, but actions perceived as escalatory often provoke other combatants to increase their own efforts, either to punish the earlier escalation or to counter its advantages. Left unchecked, cycles of provocation and counter-provocation can intensify until the cost each combatant incurs exceeds the value of its original stakes in the conflict.⁷

Escalation is not a new phenomenon, or even one particular to the modern age, but systematic thought about how to manage it did not crystallize until the Cold War, when air power, missiles, and especially nuclear weapons greatly increased the danger that any war between the East and West might quickly result in a catastrophic outcome even if leaders sought to control it. Starting in the late 1950s, growing concern about that danger inspired scholars and security analysts to examine the problem in depth and develop concepts for dealing with it. This section of the paper reviews the most prominent escalation management approaches proposed during the Cold War and explains why policymakers avoided them then and why they are even less workable today.

The Fundamental Question: Can Nuclear War Be Limited?

Early in his first term in office, U.S. President Dwight D. Eisenhower found himself on the horns of a dilemma regarding the defense of Western Europe. The NATO allies could not afford to generate enough conventional forces to offset the Soviet buildup in Eastern Europe, and Eisenhower, a fiscal conservative, was convinced that the United States’ long-term security depended on limiting spending on Defense and other government programs, thereby freeing national resources for economic development. He met this challenge by crafting a national security policy dubbed the “New Look”, which relied on strategic nuclear weapons to deter conventional and nuclear threats from the Eastern Bloc. In January 1954 Secretary of State John Foster Dulles unveiled this new policy in a speech

⁷ Morgan *et al.*, *Dangerous Thresholds*, *op. cit.*, p. 1.

in which he said, "Local defenses must be reinforced by the further deterrent of massive retaliatory power," and in order to deter aggression, the free community would have to be "willing and able to respond vigorously at places and with means of its own choosing."⁸ Although Dulles did not make the threat explicit, the implication was clear: the United States would answer any Soviet attack with nuclear *massive retaliation*.

Almost immediately, this policy came under fire. A threat of massive retaliation might be credible in response to a nuclear attack, but would Moscow believe the United States would really respond to a conventional invasion with nuclear weapons, knowing that it too had nuclear weapons with which to answer such an escalation? More seriously, how could the United States make a threat of massive retaliation credible in response to minor provocations? These questions were raised in a series of scholarly books and articles published in the mid 1950s, which proposed that strategies incorporating measured reprisals would have to be developed and limited nuclear war contemplated to make deterrent threats credible.⁹ Yet these arguments only raised more doubts. If the United States used limited nuclear strikes to blunt a Warsaw Pact invasion of Western Europe, why would Moscow not respond with its own limited nuclear strikes to return the advantage to its superior conventional forces? Would the United States then escalate its nuclear strikes and risk an even greater escalation from the Soviets? Would they be able to stop escalation short of global nuclear war? Given these considerations, would limited nuclear war even be possible? And would not even posturing nuclear forces in Western Europe in a crisis risk triggering a Soviet preemptive strike?¹⁰

Thus scholars and strategists grew increasingly skeptical regarding the feasibility of limited nuclear war. Yet one outspoken individual did believe that limited nuclear threats could be made credible, nuclear wars could be fought and won, and escalation could be controlled. That person was Herman Kahn. Kahn was the first researcher to seriously explore the strategic options available to the United States in a nuclear war and their possible effects. Using systems analysis and mathematical and scientific tools to forecast the outcomes of a series of extreme threat scenarios, he

⁸ John Foster Dulles, *Massive Retaliation*, Speech to the Council on Foreign Relations, January 12, 1954.

⁹ See for instance: William W. Kaufmann, *The Requirements of Deterrence*, Princeton, Center of International Studies, Princeton University, 1954; Bernard Brodie, "Unlimited Weapons and Limited War", *The Reporter*, Vol. 11, No. 9, November 18, 1954; William W. Kaufmann (ed.), *Military Policy and National Security*, Princeton, Princeton University Press, 1956; Henry A. Kissinger, *Nuclear Weapons and Foreign Policy*, New York, Harper and Row, 1957; Robert E. Osgood, *Limited War. The Challenge to American Security*, Chicago, University of Chicago Press, 1957; Morton Kaplan, "The Calculus of Deterrence", *World Politics*, Vol. 11, No. 1, October 1958, pp. 20-43.

¹⁰ Albert J. Wohlstetter, *The Delicate Balance of Terror*, Santa Monica, RAND Corporation, 1958; Bernard Brodie, *Strategy in the Missile Age*, Princeton, Princeton University Press, 1959, pp. 335-357; Henry A. Kissinger, *The Necessity for Choice. Prospects of American Foreign Policy*, New York, Harper, 1961. For a fuller analysis of the evolution of the Cold War-era limited war debate, see Robert E. Osgood, *Limited War Revisited*, Boulder, Westview, 1979.

concluded that the United States would need a mixed strategy that would include enough first-strike capability to reduce casualties should war appear inevitable and enough survivable retaliatory capability to make a Soviet first strike unattractive.¹¹ In subsequent work he examined the strengths and risks of 14 alternative strategies, ranging from the renunciation of war at one extreme to launching a preventive war at the other, in an exercise aimed at thinking about ways to prevent war, as well as how to "fight, survive, and terminate a war, should it occur."¹² Nevertheless, he remained bedeviled by critics who challenged the assumption that escalation could be controlled in such extreme circumstances. In 1965 he met those challenges head on with the publication of a book entitled *On Escalation: Metaphors and Scenarios*.¹³

Herman Kahn and Escalation Dominance

Kahn proposed that the United States could keep wars limited by achieving what he called *escalation dominance*. To explain his concept, he described escalation in terms of a metaphorical "escalation ladder" with each rung representing a different level of intensity in the confrontation or conflict. The lowest escalatory "rung" of Kahn's ladder represented the onset of a crisis, with higher rungs corresponding in turn to shows of force, limited conventional conflict, full-blown conventional war, limited nuclear warfare, and, at the top of the ladder, an all-out strategic nuclear exchange. Kahn acknowledged that in an actual conflict the ladder might include many more levels of escalation. In fact the notional ladder around which he organized his book was comprised of no fewer than 44 rungs, more than half of which involved at least some use of nuclear weapons. It could have been much larger, as it included few rungs involving purely conventional uses of force, and none featuring the use of chemical or biological without nuclear weapons.¹⁴

In Kahn's conception, escalation dominance describes "a capacity, other things being equal, to enable the side possessing it to enjoy marked advantages in a given region of the escalation ladder."¹⁵ It is a condition in which one has the ability to escalate a conflict in ways that would be disadvantageous or costly to the enemy, while the enemy could not do the same in return, either because it has no escalation options or because those available to it would not improve its situation. Once enemy leaders realized one had achieved escalation dominance, they should be deterred from taking the conflict to a higher rung where they would suffer greater costs with no comparable advantage. In fact, once escalation dominance is

¹¹ Herman Kahn, *On Thermonuclear War*, Princeton, Princeton University Press, 1961, p. 39. In interviews that followed the publication of this book, Kahn was notorious for declaring, "Nuclear war is winnable!"

¹² Herman Kahn, *Thinking about the Unthinkable*, New York: Horizon, 1962, p. 19.

¹³ Herman Kahn, *On Escalation. Metaphors and Scenarios*, New York, Praeger, 1965.

¹⁴ *Ibid.*, p. 39.

¹⁵ *Ibidem.*, p. 290.

achieved, the threat of further escalation should become a particularly powerful coercive lever for bringing the enemy to favorable terms.¹⁶

While Kahn's approach to escalation management is perfectly rational in an abstract context, it suffers from several serious defects when one attempts to apply it in real-world strategy making. First, the escalation ladder metaphor bears only a passing resemblance to the dynamics of actual conflict. It suggests that escalation occurs in discrete steps, observable to both sides; that the belligerents share a common perception of where each of them is standing on the ladder at any given time; and that they have sufficient control of their forces to move up or down the ladder at will. Anyone who has seriously studied crisis and war knows that such assumptions are unrealistic. Wars such as those in Korea and Vietnam illustrate how difficult it can be for adversaries to understand at what levels of conflict their opponents are attempting to fight and whether lulls or flare-ups in intensity are deliberate or circumstantial. In fact, confrontations between dangerous states are fraught with uncertainty, misperception, and miscalculation.¹⁷ Clausewitz recognized the inherent uncertainty in war and devoted considerable attention to it, discussing how incomplete, misleading, and contradictory intelligence leads to confusion and contributes to friction.¹⁸ Indeed, the "fog and friction of war" is now spoken of so commonly that the phrase has become almost cliché. In truth, war is a very uncertain affair in which only limited knowledge even of one's own forces is available at any given time. Battle management is always a challenge. In the heat of combat one's forces frequently do not do what is expected of them, either because communications have broken down, plans have disintegrated in the face of enemy resistance, or subordinate commanders have seized upon unexpected opportunities (or have simply chosen not to do what was planned).¹⁹ With so much uncertainty even regarding what

¹⁶ Daniel Byman and Matthew Waxman, *The Dynamics of Coercion. American Foreign Policy and the Limits of American Might*, Cambridge, Cambridge University Press, 2002, p. 40. Also see Daniel L. Byman, Matthew C. Waxman and Eric Larson, *Air Power as a Coercive Instrument*, Santa Monica, RAND Corporation, 1999, pp. 30-36.

¹⁷ A great deal of research was done during the Cold War on the risks of misperception. Some of the more notable works of that era include: Fred Iklé, "Can Nuclear Deterrence Last Out the Century?", *Foreign Affairs*, Vol. 51, No. 2, January 1973, pp. 267-285; Robert Jervis, *Perception and Misperception in International Politics*, Princeton, Princeton University Press, 1976; Robert Jervis, "Deterrence and Perception", *International Security*, Vol. 7, No. 3, Winter 1982-1983, pp. 3-30; Richard N. Lebow, *Between Peace and War. The Nature of International Crises*, Baltimore, The Johns Hopkins University Press, 1984, pp. 101-228.

¹⁸ Clausewitz, *On War*, *op. cit.*, pp. 117-121. Also see page 140, where he says the "general unreliability of all information is a special problem in war" of such prominence that he names it as one of three principal attributes of all military activity.

¹⁹ A classic example of several of these elements coming to play at once can be found in the opening days of World War I when the German General Staff lost contact with its western-most armies as they attempted to envelop the Allied armies in Belgium. When commander of the German First Army Alexander von Kluck saw what he believed to be an opportunity to entrap the British Expeditionary Force before it could withdraw below the Marne, he veered to the southeast,

one's own forces are doing, how much more difficult would it be to accurately determine at what "rung" an opponent is attempting to scale its efforts, especially once some number of nuclear detonations have occurred?

Second, the ladder metaphor suggests that escalation occurs along but a single dimension, vertical – i.e., increases in the *intensity* of conflict – and that it takes a conscious effort to step up to each new rung. In fact, escalation can occur along multiple dimensions, as wars in the Balkans have demonstrated, and opponents often escalate over the course of a conflict without meaning to and sometimes without even realizing they have done so. During the U.S.-Vietnam War, when U.S. and South Vietnamese forces attempted to eliminate communist sanctuaries in Cambodia and Laos, they drove their adversaries ever deeper into those countries, escalating the conflict horizontally in a way that ultimately contributed to the destabilization of the governments there. Meanwhile, communist leaders used those sanctuaries and employed other tactics to deliberately prolong the struggle in recognition that the asymmetry in stakes – defeating the insurgency was not nearly as important to the United States as expelling the Western powers and reunifying Vietnam under Hanoi's governance was to North Vietnamese and National Liberation Front leaders – gave them an advantage in motivation that would prove decisive over time.²⁰ As Fred Iklé points out, the escalation ladder metaphor fails to address this important consideration, the ability to escalate in the temporal dimension. Ultimately, contrary to what the escalation ladder metaphor implies, it often takes a greater effort to de-escalate a conflict than to escalate one, and the damage caused by escalation often cannot be undone.

Finally, true escalation dominance is rarely attainable in any challenging confrontation. It clearly was not achievable in any meaningful way during the Cold War, although both sides attempted to build arsenals at various stages of the competition that they hoped would grant them that advantage in the event of war. Even given the dramatic asymmetries of power between states in the post-Cold War world, most enemies will have some ability to escalate. Though the options may not be very attractive once the potential costs are taken into account, an adversary who finds its back against a wall often becomes remarkably inventive in discovering new ways to prolong the contest and inflict costs on the opponent in hopes of eroding its will over time. Therefore, while escalation dominance is always desirable, it is more useful to treat it as a philosophical aspiration than as a concrete policy objective.

deviating from the operational plan and unknowingly exposing his army's flank to the French Sixth Army forming up north of Paris. Only discovery of this vulnerability at the eleventh hour enabled Kluck to pull his forces back in time to avert disaster. See Larry H. Addington, *The Patterns of War Since the Eighteenth Century, Second Edition*, Bloomington, Indiana University Press, 1994, p. 140; Barbara W. Tuchman, *The Guns of August*, New York, Random House, 1962, Ballantine edition, 1992, pp. 470-491.

²⁰ Fred Charles Iklé, *Every War Must End, Second Revised Edition*, New York, Columbia University Press, 1991, p. 41.

In all fairness to Kahn, even he admitted that the escalation ladder metaphor was far from perfect.²¹ Yet, probably due to its easy visualization, it became the predominant way in which escalation was described throughout the remainder of the Cold War, and the escalation dominance approach emerged as one of two principal schools of thought about escalation management. The other school, developed by Thomas Schelling, came to be called *brinkmanship*.

Thomas Schelling and Brinkmanship

Thomas Schelling, a noted economist, was a pioneer in the use of game theory for strategy development during the Cold War.²² Drawing from that field of research, he proposed that crises and conflicts between nuclear-armed adversaries were actually contests of coercive diplomacy in which tacit bargaining was a central feature. Because neither opponent could achieve victory at an affordable cost should the contest turn into a nuclear conflagration, both shared a common interest in keeping the confrontation below the nuclear threshold. That shared interest provided a space in which they could engage in coercive bargaining, each using threats and limited applications of force to pursue its objectives at the other's expense.²³

Unlike Kahn, who made insufficient allowance for the possibilities of misperception and lack of control, Schelling made uncertainty a virtue. He argued that since nuclear-armed opponents shared the risk of escalation, one could manipulate that risk to his advantage by demonstrating that he was willing to escalate the conflict in a way that might get out of control if the adversary did not comply with coercive demands. The opponent most committed to taking the confrontation to the brink of nuclear war by binding himself to irreversible action and using "the threat that leaves something to chance" would win this contest of brinkmanship – or as Schelling sometimes described it, "game of 'chicken'" – by forcing the adversary to back down to avoid catastrophe.²⁴

²¹ In fact, he devoted an entire chapter to "Defects in the Escalation Ladder Metaphor", in which he discussed such issues as discontinuities in the importance of rungs and the spacing between them; the fact that Soviet leaders would not likely envision the same ladder or put the same importance on certain rungs as U.S. leaders; and that the concept might put undue faith in each side's rationality, clarity of understanding, and ability to communicate, particularly in conflicts at higher rungs. See Kahn, *On Escalation, op. cit.*, pp. 214-229.

²² In 2005 he and Robert J. Aumann were jointly awarded the Nobel Prize in Economic Sciences for their Cold War-era work in game theory.

²³ Thomas C. Schelling, *The Strategy of Conflict*, Cambridge, Harvard University Press, 1960, pp. 53-80; Thomas C. Schelling, *Arms and Influence*, New Haven, Yale University Press, 1966, pp. 1-34, 131-141.

²⁴ *Ibid.*, pp. 92-125. In his class at the University of Maryland in the mid 1990s, Schelling was fond of posing and answering the following question: "How do you win a game of chicken on the highway? When your car and the opponent's come careening towards each other, pull off your steering wheel and throw it out the window! *Just make sure the opponent sees you do it.*" Although Schelling is most famous for developing the brinkmanship concept, Kahn addressed it and even used the highway-chicken game metaphor, complete with the steering-wheel

Examples of brinkmanship at work can be seen in the 1948 Berlin crisis and especially the 1962 Cuban missile crisis. In the latter case, President John Kennedy and Premier Nikita Khrushchev postured nuclear and conventional forces and exchanged several letters, each warning the other that a confrontation between those forces might result in events getting beyond their control. When, after several tense days, U.S. Attorney General Robert Kennedy met privately with Soviet Ambassador Anatoly Dobrynin and warned him that the President might not be able to hold back a U.S. attack on Cuba if the crisis did not end very soon, Khrushchev lost his nerve and backed away from the brink.²⁵ The Cuban missile crisis is frequently extolled as an example of successful crisis management due to President Kennedy's skilled use of brinkmanship, and rightfully so. Yet there are several reasons why brinkmanship was not used in subsequent crises and will probably be avoided in future conflicts.

While Schelling's approach was more realistic than Kahn's in that it acknowledged the uncertainties present in confrontations between states, brinkmanship shares some of escalation dominance's defects and also exhibits other shortcomings. Like Kahn, Schelling envisioned the dynamics of escalation mostly in one-dimensional terms and occurring in a contest between two opponents relatively symmetrical in capability. Neither theorist should be condemned for such assumptions. Bipolarity was the prevailing condition of the Cold War era in which they worked, and the specter of nuclear war made other escalation risks so pale in comparison that it is not surprising that they were overlooked. Nonetheless, such shortcomings limit these concepts' utility in a world in which multiple potential adversaries, widely disparate in power, are emerging. But brinkmanship is limited even as a means of managing vertical escalation against a single opponent. Envisioned principally in terms of managing confrontations approaching the nuclear threshold, it provides little guidance for managing escalation in conflicts well below that threshold or those that have moved above it. In essence, to employ brinkmanship as a means of escalation management, one would have to deliberately take the confrontation or conflict to the brink of nuclear war.

The Cold War Escalation Management Strategy of Choice: Conflict Avoidance

Probably due to this last limitation, brinkmanship was never employed as an approach for escalation management or even crisis management after

removal strategy, in the introduction to his 1965 book, *On Escalation*. See Kahn, *On Escalation, op. cit.*, pp. 7-11.

²⁵ Graham Allison and Philip Zelikow, *Essence of Decision. Explaining the Cuban Missile Crisis, Second Edition*, New York, Longman, 1999, p. 360; Laurence Chang and Peter Kornbluh (eds.), *The Cuban Missile Crisis, 1962*, New York, The New Press, 1998, p. 378. In Robert Kennedy's memoirs he said he delivered an ultimatum to Dobrynin in that meeting, a claim that Dobrynin later denied. Nevertheless, Dobrynin did admit that, based on intelligence available to him on U.S. military preparations, he believed an airstrike or even an invasion was "very likely in the coming days." See Alexander L. George, "The Cuban Missile Crisis" in Alexander L. George (ed.), *Avoiding War. Problems of Crisis Management*, Boulder, Westview Press, 1991, p. 252.

the Cuban missile crisis. The dangers encountered in that episode so frightened U.S. and Soviet leaders that they were disinclined to further engage in such high-stakes games of chicken by any name. Military leaders on both sides continued planning for conventional and nuclear war throughout the remainder of the Cold War, but political leaders had little interest in engaging in direct confrontations, much less issuing threats that might leave something to chance. Instead, the predominant means of escalation management employed by both sides during the remainder of the Cold War became the avoidance of direct superpower conflict. The United States and Soviet Union fought a number of proxy wars between 1965 and 1990 – in Southeast Asia, Southwest Asia, Africa, and Latin America – but both were careful to avoid situations in which their own military forces might be pitted directly against each other.

One might suggest that an important lesson lies in that history, that as conflict avoidance constitutes the safest, most reliable means of escalation management, Western leaders ought to embrace it as the strategy of choice for the current age. But that would be a mistake. Conflict avoidance worked for the superpowers in the stability of a bipolar world, one in which Moscow and Washington had other actors available to posture against each other as pawns in a larger chess game. But that does not describe today's world. With the breakup of the Soviet Union, the dissolution of the Warsaw Pact, and a substantial loss of influence in the developing world, the Russian Federation has no more pawns to move and is now struggling to find its own place in the international order. Recent efforts to regain some measure of its former power have challenged Western interests in Russia's "near abroad", and Moscow may assert itself more aggressively there in the future. To employ conflict avoidance as a means of escalation management would mean abandoning Western interests in the face of those challenges. That might be an acceptable solution in cases in which Western stakes are small. It probably would not be, however, were Russia to threaten the independence or safety of states to which the West has made security commitments. NATO would then need to find a new approach to escalation management.

A New Approach to Escalation Management

Given changes in the 21st century geopolitical environment and the limitations identified in Cold War-era approaches to escalation management, a new framework for managing escalation is clearly needed, one that is both palatable to policy makers and conducive to strategy making. Such a framework should lead to a deeper understanding of the phenomenon of escalation: its dynamics, the mechanisms through which it manifests, and the motives that drive it. This section examines those elements and offers an approach for managing escalation drawn from a framework based on relationships between them. It focuses on the nature of escalation thresholds and shows how three mechanisms, working independently or in combination, can drive conflict over those thresholds. It then offers approaches for managing the mechanisms of escalation in ways that enable combatants to attain military and political objectives in war while keeping the conflict at an acceptable level of violence. Yet, despite how skillful political and military leaders are in managing escalation mechanisms, they will be working at a distinct disadvantage if their state's stakes in the issue in question are notably inferior to the opponent's. Therefore, this section concludes with the admonition that all threshold management planning – indeed, *all* strategic planning for war – should begin with an objective assessment of the balance of interests between the opponents. Leaders should then tailor their objectives and modulate their efforts accordingly.

A Closer Look at Escalation

A first step in developing a deeper understanding of escalation is to define it a way that is more precise and analytically useful. In that regard, escalation can 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.*”²⁶ Contrary to what the Cold War metaphors suggest, conflicts can intensify or expand in many ways. Some prominent examples include attacking types of targets previously considered to be off limits, opening new theaters of operations against an enemy, or employing weapons not previously used in the conflict. Further, when one examines historical cases, other less frequently considered forms of escalation emerge, such as the expansion of military objectives, the enlargement of political demands, and even increases in the vehemence of political

²⁶ Morgan *et al.*, *Dangerous Thresholds*, *op. cit.*, p. 8.

rhetoric.²⁷ Yet not every rise in threat or increase in the intensity or breadth of conflict is escalatory. Escalation only occurs when at least one of the belligerents believes that the new development has introduced a qualitative change in the crisis or conflict and behaves accordingly.

Escalation usually manifests as an interactive process between two or more opponents, each increasing its threats or use of force in response to the actions of others. But it can be unilateral as well, with one belligerent escalating to gain advantage or increase its pressure on another, independent of that actor's behavior. In such cases, the enemy might not respond because it does not have a comparable avenue of escalation, as, for example, when the United States began firebombing Japanese cities in 1945. Alternatively, an opponent might consciously choose not to answer a provocation, as was the case when Iraq began firing SCUD missiles at Israel in efforts to bring that country into the first Gulf War, or it may choose to escalate in a different way. But generally, when one actor in a conflict violates an escalatory threshold, it is reasonable to expect its enemies to follow suit. A threshold breached tends to lose its saliency. Yet even this is not always the case, particularly in conflicts between adversaries who are markedly dissimilar in capability. For several years during the Vietnam War communist forces violated the neutrality of Cambodia and Laos to move supplies, train and reconstitute forces, and launch operations into South Vietnam before U.S. and South Vietnamese forces launched major operations into those states. Yet the international community and even the U.S. public viewed the latter actions, when they finally did occur, as a significant escalation.

Escalation in armed conflict is a very diverse phenomenon. It can occur quickly or slowly. A belligerent can escalate in dramatic moves that are visible to almost any observer, or in incremental steps so small that they are unrecognized as constituting significant escalation until after the fact, even by the one doing the escalation. Given this diversity, to recognize escalation and understand it more fully, we must examine the nature of thresholds more closely.

The Nature of Escalation Thresholds

An escalation threshold is an identifiable point in the intensity or scope of events which when crossed is recognized by at least one of the belligerents as constituting a significant change in the nature of the conflict. Thresholds are socially constructed elements existing purely in the minds of the parties involved, so they come in many forms. Some escalation thresholds are symmetrical in that all parties to a conflict recognize them and tend to view them similarly. Examples might include being the first to initiate hostilities in a crisis or employ nuclear weapons in a war. But sometimes thresholds that are important to one actor may seem trivial or even be invisible to another.

²⁷ One might not think of the last example as constituting a dangerous form of escalation. But consider how, on the eve of the first Gulf War, when Saddam Hussein indicated a willingness to reach a negotiated settlement, President George H.W. Bush escalated his public rhetoric making it impossible for the Iraqi leader to back down without losing face in the Arab world.

This subjectivity is one of the reasons why escalation can be difficult to recognize, control, manage, and exploit.²⁸

The subjective nature of thresholds creates serious risks of misperception and miscalculation. If one party knows that another considers a particular threshold to be important, that threshold is likely to be significant in its own eyes as well. But the adversary's perspective is not always well known or understood, nor is it always clear whether the enemy knows where one's own thresholds lie or what importance one places on those thresholds. In general, the thresholds that will be the easiest to anticipate are those that are geographically prominent – such as a river recognized as the boundary of one's territory – or those involving strongly held international norms, such as the taboos against the use of nuclear, chemical, and biological weapons. National policies firmly established before a crisis or conflict arises can also offer signposts regarding what actions an opponent might consider escalatory in war. Escalation thresholds might be relevant to a wide range of adversaries and in a broad set of circumstances, such as the use of nuclear weapons, or they might be particular to specific cases, such as the 19th century agreement between the European great powers guaranteeing Belgium's independence and neutrality. Thresholds thought firm in peacetime may be viewed differently when their violation actually occurs. An international ban on unrestricted submarine warfare, widely supported in the interwar years, quickly dissolved after the outbreak of World War II. Yet assuming that thresholds have grown weak with age can be dangerous, as German leaders discovered when they violated Belgian neutrality in 1914, thinking Britain would surely not go to war over a mere “scrap of paper”.²⁹

Given the subjective nature of thresholds, states sometimes attempt to manipulate them to their own advantage, either to strengthen or create new thresholds to better deter an enemy from undertaking an undesired action, or to reduce the significance of established thresholds to make crossing them less risky. Achieving the former objective may involve employing exaggerated rhetoric to demonize the use of certain weapons, or more concerted political approaches such as formally outlawing them.³⁰ Strengthening a threshold in peacetime is challenging in that it requires building an international consensus on an issue that would likely advantage some states at the expense of others, but persuading members of the international community that a threshold is less important than they previously believed is even more difficult. States hold thresholds to be important because their violation puts them at greater risk.

²⁸ See Schelling, *Arms and Influence*, *op. cit.*, pp. 153-168, 283-286.

²⁹ This was part of German Chancellor Theobald von Bethmann-Hollweg's indignant response when Ambassador Sir Edward Goschen delivered Britain's ultimatum for Germany to cease hostilities against Belgium within 48 hours. See Tuchman, *The Guns of August*, *op. cit.*, pp. 153-154.

³⁰ The Chemical Weapons Convention exemplifies an effort to strengthen an established threshold by formally outlawing a class of weapons. Alternatively, declaring that chemical, biological, and nuclear weapons collectively constitute a single category of “weapons of mass destruction”, despite the obvious differences in their destructive power, is a prominent example of efforts to raise a threshold via demonization.

Thresholds that emerge during a conflict are the most difficult to anticipate, as they typically arise in response to events the possibility of which are unforeseen before they occur, or capabilities the threatening nature of which are unappreciated before they are employed. Post-Cold War conflicts have revealed that the vast conventional superiority that Western forces now enjoy increases chances that opponents facing those capabilities will perceive some actions escalatory that Western military leaders consider routine. For example, early in NATO's 1995 air campaign against the Bosnian Serb Army, an insufficient number of Allied strike aircraft were available to service targets across all of Bosnia, so Allied planners confined their efforts to the country's southeast zone of operations (ZOA). When additional aircraft arrived in theater, the Allies wanted to engage a greater number of targets, but Serbian air defenses in the northwest ZOA were particularly dense, so planners decided to employ F-117A stealth fighters and U.S. Navy Tomahawk Land-Attack Missiles (TLAMs) to "soften them up" before sending in non-stealthy strike aircraft. That decision was based purely on these weapons' availability and operational utility, yet Serbian leaders interpreted their use as a major escalation in NATO's prosecution of the war.³¹

Understanding Mechanisms and Motives

While attention to thresholds is an important prerequisite for managing escalation, that alone is not enough. Leaders must also understand the mechanisms through which escalation manifests and why parties to a conflict sometimes choose to intensify or broaden the scope of their attacks, even while hoping to keep the conflict limited. Escalation in confrontation and war occurs through three mechanisms: *deliberate*, *inadvertent*, and *accidental*. While these mechanisms are theoretically distinct, escalation in an actual conflict can result from the interaction of more than one of them at once, and escalation of one type can sometimes trigger escalation through one or both of the other mechanisms as well.

Deliberate Escalation

Deliberate escalation occurs when a party to a confrontation or conflict intentionally undertakes some action that it knows will cross one or more of an opponent's escalation thresholds. There might be any number of proximate motives for taking such action, but they can all be generally described as either *instrumental* or *suggestive* in nature or some combination of both. In instrumentally motivated escalation an actor believes that increasing the intensity or scope of the fight will work to its advantage by raising its prospects of success. A belligerent might throw in an extra division to turn the tide in a land battle,³² broaden the list of

³¹ Mark J. Conversino, "Executing Deliberate Force, 30 August-14 September 1995", in Robert C. Owen (ed.), *Deliberate Force. A Case Study in Effective Air Campaigning*, Maxwell AFB, Air University Press, 2000, pp. 150-153. Also see Richard L. Sargent, "Weapons Used in Deliberate Force", in Owen (ed.), *Deliberate Force, op. cit.*, p. 264.

³² Committing additional units to a battle is not necessarily escalatory. If a conflict has already escalated to the point at which each side expects the other to do whatever it possibly can to win at the conventional level of war, then neither is surprised when the other commits forces previously held in operational or strategic

bombing targets to overwhelm an enemy's capacity to resist, or launch an attack in a region previously unthreatened to cut off the enemy's access to important resources or force it to divide its forces. Escalation of this type often prompts the opponent to try to match or surpass the increase in effort, or escalate in some other dimension, to counter its perceived advantage. Alternatively, in suggestively motivated escalation a belligerent deliberately increases the intensity or scope of conflict in efforts to signal an opponent that it ought to change its behavior in some way. This form of deliberate escalation is akin to the kind of coercive bargaining that Schelling described, in which an actor punishes its opponent, not primarily for the direct military benefit that might result from such action, but to suggest that more punishment will come if the opponent does not comply with coercive demands.³³

Operation Rolling Thunder, the U.S. bombing campaign against North Vietnam between 1965 and 1968, is probably the most frequently mentioned example of this kind of coercive escalation. Critics often cite the failure of Rolling Thunder as condemnation of strategies that use "graduated escalation" for purposes of signaling or otherwise restrain the employment of air power.³⁴ Their arguments have merit to the extent that Rolling Thunder was indeed overly restrained, at least in the early phases. However, it is important to understand that using deliberate escalation for coercive signaling does not imply that attacks need to be gradual or excessively restrained. The systematic firebombing of Japanese cities in 1945 was an example of deliberate escalation for suggestive motives. It signaled Japanese leaders that until they complied with Allied demands for unconditional surrender, they could expect such horrendous costs to mount, day after day. The atomic bombings of Hiroshima and Nagasaki constituted another deliberate and dramatic escalation to reinforce that signal.³⁵

reserve. However, if both sides have previously withheld the employment of forces below a given threshold in hopes of keeping the conflict limited, then one opponent commits additional forces exceeding that threshold, the opponent is likely to consider the move escalatory and react accordingly. In sum, the commitment of additional force is escalatory if and only if either of the opponents believes doing so has crossed an escalation threshold.

³³ Schelling, *Arms and Influence*, *op. cit.*, p. 172.

³⁴ See for instance A. L. Gropman, "The Air War in Vietnam, 1961-73", in R. A. Mason (ed.), *War in the Third Dimension. Essays in Contemporary Air Power*, London, Brassey's Defence Publishers, 1986, pp. 37-39. Robert Pape also condemns Operation Rolling Thunder, but the root of his criticism is not that the strategy was overly restrained. Rather, he argues that North Vietnam was largely immune to conventional coercion during that period because conventional bombing was ineffective in defeating a guerilla warfare strategy and Hanoi was willing to bear whatever costs the United States was willing to inflict to achieve its territorial ambitions in South Vietnam. See Robert A. Pape, *Bombing to Win. Air Power and Coercion in War*, Ithaca, Cornell University Press, 1996, pp. 174-195. Also see Robert A. Pape, "Coercive Air Power in the Vietnam War", *International Security*, Vol. 15, No. 2, Fall 1990, pp. 103-146.

³⁵ It is important to note that, while the atomic bombings represented a dramatic escalation in that they introduced a terrifying new weapon, they were not escalatory in terms of the levels of destruction or suffering they caused as

Avoiding deliberate escalation is partly a matter of self-restraint. Leaders should resist the temptation to escalate a conflict in ways that might offer temporary tactical advantages at risk of suffering serious long-term strategic costs. But self-restraint in war can be exceedingly difficult. The more restrained one is, the more difficult it is to achieve one's military objectives. Efforts to exercise restraint often pit military leaders, who tend to argue for more operational freedom, against their political superiors, who worry that granting such freedom might result in a more intense and costly conflict or bring other belligerents into the war. In the 1962 Cuban missile crisis, for instance, U.S. military leaders lobbied for an invasion of Cuba or, at least, air strikes against the Soviet offensive missiles being installed there, but President Kennedy wisely resisted those pressures.³⁶ Similarly, the U.S. experience in the Korean War, when General Douglas MacArthur's 1950 penetration into North Korea prompted Chinese intervention and a costly three-year war of attrition, made President Lyndon Johnson resistant to U.S. military pleas for permission to launch a more intense bombing campaign against North Vietnam in 1965.³⁷

But containing deliberate escalation requires more than just self-restraint. As war is a struggle between two or more adversaries, one must also deter other actors from escalating the conflict by convincing them that doing so would not work to their advantage. Deterrence involves threatening to punish the opponent for some prospective escalation, orchestrating forces in a way that convinces it that the escalation would not be successful, or some combination of both approaches. The objective is to influence the opponent's decision calculus, leading enemy leaders to conclude that the costs of escalation would ultimately outweigh whatever benefit they might hope to gain from it. Deterring deliberate escalation often involves threats of counter-escalation. In some relatively easy cases, simply threatening to match the escalation symmetrically might be enough to deter it. The presence of a powerful bomber force might be enough to deter an adversary from embarking on a campaign of city bombing. Similarly, the vulnerability of an enemy's heretofore-unmolested province might be enough to deter it from escalating a conflict into a region in one's own country not yet affected by the war. In more challenging cases, however, greater or different threats might be required to offset the advantages the enemy expects to gain by escalating. France's nuclear

compared with the firebombings. For instance, the March 9, 1945, firebombing of Tokyo inflicted 185,000 casualties in a single attack. In comparison, the combined casualty toll from the bombing of Hiroshima and Nagasaki was about 110,000 people killed with an estimated 90,000 more people injured. See *United States Strategic Bombing Survey: Summary Report (Pacific War)*, Washington, 1946, reprinted as *The United States Strategic Bombing Surveys (European War)*, (Pacific War), Maxwell AFB, Air University Press, 1987, pp. 92, 100-101.

³⁶ We now know that some of the missiles became operational during the crisis, were armed with nuclear warheads, and launch authority had been delegated to the Soviet military commander in Cuba. Graham Allison and Philip Zelikow, *Essence of Decision*, *op. cit.*, pp. 215-217; Lyle Goldstein, *Preventive Attack and Weapons of Mass Destruction. A Comparative Historical Analysis*, Stanford, Stanford University Press, 2006, p. 45.

³⁷ Mark Clodfelter, *The Limits of Airpower. The American Bombing of North Vietnam*, New York, The Free Press, 1989, p. 53; David Halberstam, *The Best and the Brightest*, New York, Fawcett Crest, 1972, p. 424.

doctrine during the Cold War, “deterrence of the strong by the weak”, offers an example of such an asymmetry in its reliance on the imbalance of interests favoring the defender to deter the aggressive designs of more powerful potential adversaries.

Whether instrumental or suggestive, deliberate escalation is the mechanism most naturally associated with the metaphor of climbing a ladder. Therefore, it is what military and political decision makers tend to envision when they think of escalation, leading them to assume they can control it. Unfortunately, escalation often gets out of control despite the best efforts of leaders on all sides to contain it. This is because not all escalation is deliberate in nature. Sometimes it occurs inadvertently or due to accident.

Inadvertent Escalation

Inadvertent escalation occurs when one belligerent deliberately undertakes an action that it does not consider escalatory, but the action is perceived as such by an opponent. In other words, the action crosses a threshold that is important to the adversary, but appears insignificant or is unknown to the escalator. Incidents of inadvertent escalation typically result from not anticipating how an opponent will view certain actions, either due to a lack of intelligence or simply not considering how the opponent’s view of the conflict, and particularly its vulnerabilities, will likely affect its perception of thresholds. It can also result from an inability to anticipate the reactions of third parties or other second- or third-order consequences.³⁸

Numerous cases of inadvertent escalation can be found in past wars. Several, such as Germany’s 1914 violation of Belgian neutrality, MacArthur’s 1950 drive into North Korea, and the United States’ expansion of the Vietnam War into Cambodia and Laos have already been mentioned. Among the many insights these cases offer is that inadvertent escalation cannot be directly deterred because it occurs as a result of decision makers not understanding the degree to which the actions they are embarking upon are escalatory. Therefore, a straightforward approach to reducing the risk of inadvertent escalation would be to, first, inform adversaries of where one’s important escalation thresholds lie then issue threats or take other actions to deter them from violating those thresholds. Yet this too is easier said than done. Inadvertent escalation often occurs because neither side has fully considered where even its own escalation thresholds lie until one of them is crossed. In other cases it occurs because one side considers a threshold to be so obvious that it need not warn the other side of its existence.³⁹ Complicating matters, thresholds often change over the course

³⁸ For the seminal work illuminating the risks of inadvertent escalation during the Cold War, see Barry R. Posen, “Inadvertent Nuclear War? Escalation and NATO’s Northern Flank”, *International Security*, Vol. 7, No. 2, Autumn 1982, pp. 28-54; and Barry R. Posen, *Inadvertent Escalation. Conventional War and Nuclear Risks*, Ithaca, Cornell University Press, 1991.

³⁹ This raises questions about why thresholds are sometimes perceived by one side and not the other or perceived differently by opponents. A range of causes might be involved, such as cultural differences, bureaucratic routines, and various forms of cognitive biases. The causes doubtless vary from case to case, and they

of a conflict, and belligerents may resist revealing where their critical thresholds lie, because to do so would acknowledge certain political or military vulnerabilities. A threshold illuminated provides a focal point for deterrence, but it also exposes a weakness that enemies might choose to exploit. A deliberate effort to keep one's critical thresholds vague is illustrated in France's nuclear doctrine, which declares that states that threaten its vital interests are at risk of nuclear retaliation without specifying what those vital interests are or even whether they would have to be overtly attacked before France would strike.⁴⁰

Given these challenges, managing risks of inadvertent escalation requires a balanced strategy incorporating several features. The first step is to make a considerable effort in advance to identify potential paths of escalation. This requires not only collecting and analyzing intelligence about each adversary's capabilities, vulnerabilities, and potential attitudes and behaviors, but also those of important third parties and an assessment of one's own thresholds as well. Next, analysts must sensitize planners and decision makers to the risks of inadvertent escalation, both generally and in terms of specific escalation thresholds relevant to the contingency at hand, so they can consider those elements in their planning. Finally, strategic plans need to incorporate features designed to avoid critical escalation thresholds of other actors and steer enemy actions away from one's own thresholds, either by announcing their existence and issuing threats to deter their violation, or by visibly posturing forces in ways that deter enemy exploitation by denying benefits of that behavior.

Accidental Escalation

Perhaps the most difficult form of escalation to manage directly is that which occurs totally by accident. Like inadvertent escalation, accidental escalation is unanticipated, but instead of being an unexpected result of deliberate action, it is the consequence of events that were not intended in the first place. Such events might be the results of pure accident, such as sinking a ship belonging to a neutral state due to misidentification, or bombing the wrong target due to a navigation error or outdated map. But accidental escalation can also result from military forces acting in ways not authorized or intended by national leaders, either because the combatants do not understand their leaders' intent or because they do, but disregard it and act on their own.

Twentieth century wars offer numerous examples of accidental escalation. Those of the first type – that is, escalation resulting from *pure* accident – are often exemplified by the escalation in strategic bombing that

probably operate interdependently in some cases. The author has not done sufficient comparative analysis of cases to determine which causes occur most frequently or exert the most influence, but his impression so far is that a frequent cause is simply the execution of military operations that one side considers routine, due to established doctrine and standard operating procedures, without sufficient consideration made regarding whether such actions will cross any of the opponent's critical thresholds.

⁴⁰ David S. Yost, "France's New Nuclear Doctrine", *International Affairs*, Vol. 82, No. 4, 2006, pp. 701-721.

occurred early in World War II after the Luftwaffe accidentally bombed London on August 24, 1940. According to some historians, Britain's retaliation against Berlin the following night enraged Adolf Hitler, contributing to his decision to launch the Blitz, the bombing campaign against London and other British cities that lasted into 1941.⁴¹ An often cited example of accidental escalation resulting from a combatant commander deliberately exceeding his superior's intent occurred late in the Vietnam War, when U.S. 7th Air Force commander General John D. Lavelle authorized aircrews to engage surface-to-air missile (SAM) sites in North Vietnam before those sites fired on U.S. aircraft, a violation of the standing rules of engagement (ROE), and told aircrews to report that the SAMs had fired first.⁴²

Since accidental escalation, like inadvertent escalation, is not intentional, one can do little to deter the enemy from doing it. The best leaders can hope to do is recognize that isolated incidents of enemy provocation might not be deliberate and modulate their responses to those events accordingly. That does not mean that they should ignore all provocations believed to be accidental. Failing to respond firmly to undue aggression, even when that aggression was not done deliberately, might signal a lack of resolve that emboldens even greater escalation. But military and political leaders do need to evaluate each incident in context and respond to it in a judicious manner.⁴³

⁴¹ See, for instance, Christopher Catherwood, *Winston Churchill. The Flawed Genius of World War II*, New York, Berkley, 2009, p. 72; Terry Copp, *No Price Too High. Canadians and the Second World War*, Whitby, McGraw-Hill Ryerson, 1996, p. 50. Other historians argue that Hitler would soon have ordered the bombing effort to shift from RAF airfields to London, anyway, because time was running out for executing Operation Sea Lion, the invasion of Britain, before fall weather would make a channel crossing infeasible. Breaking the resistance of RAF Fighter Command was a prerequisite to the invasion, and Luftwaffe leaders believed that only attacking London would draw British fighters up in sufficient numbers to enable Fighter Command's destruction. See Richard J. Overy, *The Air War. 1939-1945*, Chelsea, Scarborough House, 1980, pp. 34-36; John Ray, *The Battle of Britain. Dowding and the First Victory, 1940*, London, Cassell & Co., 2000, pp. 92-93.

⁴² When this practice was exposed in the news media, General Lavelle was relieved of command and forced to retire in disgrace at a reduced grade. The Lavelle case is often studied in courses on professional ethics in U.S. military colleges and universities, the dilemma being: Is a military commander justified in violating the orders of superiors when those orders put the forces under his or her command in peril for no sound military objective? Ironically, on August 6, 2010, President Barack Obama, after an extensive review of the Lavelle case, exonerated the general and posthumously restored him to his full military rank and honors. The investigation revealed that General Lavelle's conduct had been consistent with secret orders from President Richard Nixon, who had, for political reasons, elected not to speak up for him when the scandal erupted. See <http://www.arlingtoncemetery.net/jlavelle.htm>, accessed: June 19, 2011.

⁴³ An example of a peacetime accident that could have escalated into an international crisis is the April 2001 collision between a U.S. Navy P-3 Orion reconnaissance aircraft and a Chinese F-8 fighter over the South China Sea. When this incident occurred, I was a faculty member at the U.S. Air Force School of

Regarding one's own forces, military leaders need to identify those factors that might raise risks of accidental escalation in any confrontation or war and manage them. Approaches for managing these risks depend on the nature of possible accidents. Those emanating from potential mechanical failures can be reduced by designing systems in ways that allow for high degrees of reliability and by creating procedures and ROE that minimize such risks. Risks of accident arising from human error or carelessness can be minimized by training, exercise, and closer leadership attention. Risks of escalation resulting from subordinates misunderstanding commanders' intent or deliberately defying the limitations placed upon them can only be reduced by developing reasonable and coherent ROE and communicating and enforcing them more effectively. In sum, minimizing the risk of accidental escalation requires effective leadership in all phases of a military operation and at all levels in the chain of command.

Even when strong leadership and discipline are applied, however, such risks can never be completely eliminated. After all, some accidents will always happen. When they do, allied leaders will have to promptly assess their potential impacts and take whatever actions are needed to mitigate their escalatory effects. Such actions might include informing the adversary that the act was unauthorized and will not be repeated. Some accidents might also require issuing threats or posturing forces in efforts to deter the adversary from escalating in response to them. But in some cases, leaders will simply have to accept the fact that the adversary will likely escalate in response to the accident, endeavor to establish a new upper-escalation threshold, and resolve to fight on to victory, albeit at higher costs. As Clausewitz argued, friction in war is an inescapable reality. Only practice, experience, and the unrelenting will of a competent commander can overcome it.⁴⁴

The Importance of Understanding the Balance of Interests

Regardless of how skillful a nation's military and political leaders are in threshold management, they will find it difficult to obtain their objectives at affordable costs if their interests in the dispute are substantially less than the opponent's. Wars result from conflicts of interest, however defined.⁴⁵

Advanced Air and Space Studies, and several students asked me: "What is China up to? Why would they do this to us?" My response was, "What makes you think anyone did this deliberately, and if so, that Beijing was behind it, versus someone lower in the PLA chain of command or even the pilots themselves, Chinese or American?" Fortunately, military and political leaders on both sides kept cool heads, and the issue was resolved with the U.S. aircrew's return after 11 days and the aircraft about three months later. For an analysis of this episode, see Shirley A. Kan, Richard Best, Christopher Bolcom, Robert Chapman, Richard Cronin, Kerry Dumbaugh, Stuart Goldman, Mark Manyin, Wayne Morrison, Ronald O'Rourke and David Ackerman, *China-U.S. Aircraft Collision Incident of April 2001. Assessment and Policy Implications*, CRS Report for Congress, Washington, Congressional Research Service, October 10, 2001.

⁴⁴ Clausewitz, *On War*, *op. cit.*, pp. 119-121.

⁴⁵ I admit that this statement expresses an unabashedly realist viewpoint. Nevertheless, I maintain that, while there are many differences in frames of reference and worldviews between the leading international relations theories, many of the fundamental differences are more about how they define, label, and

And as Clausewitz stated, “The political object – the original motive for the war – will thus determine both the military objective to be obtained and the amount of effort it requires.”⁴⁶ As the value of the political object also determines the level of motivation to pursue that object and the resolve to carry on in the face of resistance, the relative stakes in the conflict – that is, the balance of interests between the adversaries – also influences each side’s perception of escalation thresholds and their tolerances for costs and risks.⁴⁷

Unfortunately, some of the wars in which the Western powers have engaged since World War II have demonstrated that national leaders tend to overestimate their interests, vis-à-vis those of the adversary, early in the conflict, even while underestimating the levels of effort that will be required to obtain those interests. Vietnam offers the archetypal example. In that conflict U.S. leaders initially believed that stopping the spread of communism there would be crucial to the success of the policy of containment, which was considered a core U.S. national security interest. Conversely, they assumed that North Vietnam’s interests in the dispute derived simply from an ideologically motivated ambition for conquest – and the interests of the National Liberation Front were dismissed as being subservient to Hanoi’s – failing to grasp that both actors, though communist, were primarily motivated by nationalist aspirations and anti-colonial resentment. As a result, although military superiority enabled the United States to escalate the conflict in ways that inflicted enormous costs on its opponents, the U.S. public, as the war ground on, grew disillusioned regarding U.S. interests there and increasingly intolerant of the costs the nation was paying. Ultimately, the communists enjoyed an asymmetry of interests that translated into an asymmetry of motivation, making them much more persistent and cost tolerant than U.S. citizens or their government.⁴⁸

Similar dynamics can be seen in some of the stability operations in which the West has engaged, such as those in Beirut, Lebanon (1982-1984), and Mogadishu, Somalia (1992-1994). In these episodes, Western states intervened for a variety of motives, ranging from humanitarian concerns to desires to eliminate or contain sources of instability before they further jeopardized regional economic and security interests. Yet, in both of

prioritize interests – i.e., whether they stem more from the pursuit of power, security, and wealth or from values, norms, and identities – than disagreements about whether actors fight over them. Actors fight over conflicts of interest, however defined.

⁴⁶ Clausewitz, *On War*, *op. cit.*, p. 81.

⁴⁷ Alexander L. George, “Theory and Practice”, in Alexander L. George and William E. Simons, *The Limits of Coercive Diplomacy, Second Edition*, Boulder, Westview, 1994, p. 15.

⁴⁸ See William E. Simons, “U.S. Coercive Pressure on North Vietnam”, in George and Simons, *The Limits of Coercive Diplomacy*, *op. cit.*, pp. 133-173; and Alexander L. George and William E. Simons, “Findings and Conclusions”, in George and Simons, *The Limits of Coercive Diplomacy*, *op. cit.*, pp. 281-282. Also see Wallace J. Thies, *When Governments Collide. Coercion and Diplomacy in the Vietnam Conflict, 1964-1968*, Berkeley, University of California Press, 1980, pp. 6-13.

these efforts, local actors' interests were much greater than those of the intervening powers, and they achieved escalation dominance by demonstrating their willingness and ability to inflict casualties on stability forces at levels that exceeded the cost tolerances of Western governments and publics at home.⁴⁹

Considering these examples, one might conclude that Western states will be at an inherent disadvantage in most future conflicts because they will likely be conducting expeditionary operations against opponents on their own territories. In such settings, Western forces would be attempting to wage limited wars for limited stakes against adversaries who would have much greater stakes in the outcomes and therefore feel less bound by the same constraints. That has indeed been a problem in many past conflicts, and it will likely continue to be a troubling dynamic in some future cases, particularly those in which the West intervenes in civil wars or attempts to stabilize failed states in regions in which Western economic and security interests are only marginal.

However, there is no reason to assume *a priori* that such would be the case in the kinds of future conflicts in which the West would be most concerned about managing escalation – those with nuclear-armed states. Granted, in any war against such an opponent, Western forces would likely be conducting expeditionary operations, but that, in itself, does not indicate the West would not perceive vital interests to be at stake. After all, several states in the Western Hemisphere concluded that their interests in Europe and Asia were great enough to mount considerable expeditionary operations and pay substantial costs in two world wars during the twentieth century, even though their homelands were not directly threatened. Moreover, the West might not be the only side embarking on expeditionary operations. Depending on what interests are at stake, a future nuclear-armed adversary might be sufficiently motivated and emboldened to conduct military operations outside its home territory as well. Whether a significant asymmetry of interests would emerge should the West confront such a move, and if so, which side would perceive greater stakes in the issue (and manage to sustain those perceptions in the face of costly resistance) would depend on a wide range of factors. One important consideration would be which side is trying to change the status quo and which is trying to preserve it. Beyond that, were third-party territory to be fought over, the cultural orientation and political preferences of the government and citizens of that territory would doubtless play important roles in shaping local, regional, and international perceptions regarding which side had greater moral justification in attempting to change or defend the status quo.⁵⁰

⁴⁹ For case studies examining the escalation dynamics that occurred in these operations and an analysis of the escalation dynamics of irregular warfare more generally, see Morgan *et al.*, *Dangerous Thresholds*, *op. cit.*, pp. 117-158, 197-220.

⁵⁰ For instance, consider differences in world reactions to the Anschluss, Hitler's 1938 annexation of Austria, a Germanic nation whose citizenry included an

In any event, before intervening in a regional conflict against a nuclear-armed opponent – indeed, before embarking on any military expedition in the face of potential resistance – national leaders should do a serious, dispassionate assessment of each side's interests in the issue at hand and scope their objectives accordingly. Such an assessment would be an essential element of the threshold analysis, as thresholds are closely related to perceptions of interest as well as to vulnerabilities. Once each side's stakes and potential escalation thresholds are understood, Western leaders should craft and execute an operational plan aimed at defeating the opponent's conventional forces, but only to the extent needed to obtain their limited objectives. Ideally, that plan would respect the opponent's higher-level interests and thresholds, while holding them at risk as part of a carefully crafted escalation management strategy.

Finally, as any conflict unfolds, events will inevitably deviate from expectations, and plans and operations will have to be adjusted. The fortunes of war could go either way, but leaders must maintain a steady hand regardless. They should resist the temptation to escalate their political objectives in the face of unexpected military success.⁵¹ Conversely, in the event of serious setbacks, military leaders will want to increase their efforts to avoid defeat. This is permissible – indeed, the opponent will likely expect it – so long as such increases do not violate important escalation thresholds. Should it appear that defeat is unavoidable without a serious escalation, political and military leaders should assess the implications of such an escalation in terms of the risks involved and weigh them against the interests at stake before embarking down a path on which the potential costs might ultimately exceed even the benefits of victory.

outspoken pro-Nazi, pro-unification minority, and the German annexation of Czechoslovakia and especially the invasion of Poland in 1939.

⁵¹ Iklé, "Can Nuclear Deterrence Last Out the Century?", *op. cit.*, pp. 10-11.

Managing Escalation Risks in a Conflict with the Russian Federation

Russia is not an enemy of the United States or its allies, and there is little danger that the West will find itself in armed conflict with this state in the foreseeable future. But while Moscow currently enjoys amicable relations with the West, there are serious stress points in the relationship and it is not inconceivable that a future conflict of interest between Russia and one of the states on its periphery would result in a military confrontation with NATO. Such a crisis would evoke serious risks of escalation. This section of the paper examines possible routes to a Russo-NATO confrontation and the escalation dynamics that might arise should that confrontation result in armed conflict. Then it discusses ways in which Western leaders could employ threshold management concepts to minimize the risks of escalation.

Potential Routes to Conflict between Russia and NATO

Moscow's relationship with the West is strongly colored by the decline in prestige Russians believe they have suffered as a result of the breakup of the Soviet Union and Russia's loss of superpower status. The bitterness associated with these events was worsened by the failed liberalization policies of the Yeltsin era and the sense that Yeltsin's efforts to engage the West were met with exploitation at Russia's expense. Those efforts culminated with Yeltsin's resignation on December 31, 1999, six months after NATO embarrassed Moscow by conducting combat operations against Russia's traditional client state, Serbia.⁵² Until that point, Moscow had been patient and largely docile in the face of Western expansion. From the Kosovo war onward, however, Russian leaders began to push back, as first evidenced in the Russian military occupation of Pristina International Airport over NATO objections.⁵³

⁵² In a July 2001 discussion with a professor from the University of Moscow at the Asia-Pacific Center for Security Studies, Honolulu, Hawaii, the author was made acutely aware of Russia's sense of indignation over NATO conducting military operations in its traditional sphere of influence.

⁵³ "Russian Troops Camp in Pristina", *BBC World: Europe*, June 12, 1999, available at: <http://news.bbc.co.uk/2/hi/europe/367490.stm>, accessed January 16, 2012; Patrick Wintour, Ian Traynor and Tom Whitehouse, "Russian and British Troops in Tense Pristina Standoff", *The Observer*, June 12, 1999, available at:

These experiences reinforced Russia's traditional worldview that interstate relations are inherently competitive in nature and the only reliable path to national security is the enhancement of state power.⁵⁴ Consequently, national goals under President and later Prime Minister Vladimir Putin and President Dmitry Medvedev have focused on reestablishing domestic stability and economic growth, restoring national self-esteem and international prestige, and projecting Moscow's influence on Russia's periphery, particularly among the former Soviet republics.⁵⁵

It is in this last regard that the potential seed of future conflict germinates. Many in Moscow deeply resent the fact that NATO has expanded eastward to the Russian border.⁵⁶ "There is general agreement in Russian government and analytic circles that NATO expansion threatens Russia's interests, particularly as it continues to reach deeper and deeper into what Russia sees as its own sphere of influence."⁵⁷ Russians see this development – along with Washington's criticism of Moscow's domestic policies and U.S. efforts to strengthen democracy in countries on Russia's periphery – as evidence of a broader Western campaign to undermine Russian prestige and power. Seen through this lens, Ukraine's Orange Revolution and Georgia's Rose Revolution were not popular responses to contested elections, but coups engineered in the West to replace pro-Russian governments with pro-Western ones.⁵⁸ Russians are particularly resentful for what they believed to be U.S. encouragement and support for Georgia's actions in the 2008 conflict with the breakaway provinces, Abkhazia and South Ossetia, and for what they perceived to be biased reporting of that affair in the Western media.⁵⁹

Given these perennial sources of tension, there are several flashpoints on Russia's periphery at which the United States and its allies could find themselves in confrontation with Moscow. Such a confrontation could have occurred in Georgia in 2008 had the United States chosen to support that NATO aspirant against Russian Federation forces when they intervened there. Georgia's deep resentment for Russia's intervention and Moscow's subsequent recognition of Abkhazia and South Ossetia as independent states increases risks that another confrontation might occur in the future. The aggressiveness with which Russia prosecuted its military operations⁶⁰ was meant to send a message deterring Georgia and the other

<http://www.guardian.co.uk/world/1999/jun/13/balkans5>, accessed: January 16, 2012.

⁵⁴ Dean and Harlow, *Russia Workshop, op. cit.*, pp. 4, 6, 11-13.

⁵⁵ Oliker *et al.*, *Russian Foreign Policy, op. cit.*, pp. 83-95.

⁵⁶ Dean and Harlow, *Russia Workshop, op. cit.*, p. 16.

⁵⁷ Oliker *et al.*, *Russian Foreign Policy, op. cit.*, p. 111.

⁵⁸ See for instance: "Russian Federation Military Policy in the Area of International Information Security: Regional Aspect", *Military Thought*, Vol. 16, No. 1, 2007, pp. 3-4 and European Union, *Independent International Fact-Finding Mission on the Conflict in Georgia Report* (IIFFMCG Report), Vol. II, September 2009, p. 23.

⁵⁹ Oliker *et al.*, *Russian Foreign Policy, op. cit.*, p. 126-131.

⁶⁰ While the EU's Independent International Fact-Finding Mission on the Conflict in Georgia found that Russian military intervention in South Ossetia was justified in defense of Russian peacekeepers there, "the Russian military action went far beyond the reasonable limits of defence..." in regards to the "massive and

former Soviet republics from further assertiveness, and it might indeed have that effect.⁶¹ However, it could also trigger an opposite reaction, generating political pressure on Washington and other NATO capitals to support those states in any future confrontation and making those states more strident in expectation of that support. Such a crisis could erupt in Ukraine, where tensions have periodically arisen over such issues as its relations with the Western Alliance, the terms of Russia's continued use of the naval base at Sevastopol, and Russia's efforts to use natural gas access and pricing as a coercive lever over Kiev. Even greater friction exists between Russia and NATO members Estonia and Latvia over those countries' accession to the Alliance, their alleged discrimination against Russian-speaking citizens, their insistence that borders set in the Soviet era are unfair and should be adjusted in their favor, and the general resentment they harbor towards Russia for its historical domination.⁶²

Even global warming spawns new risks of conflict. As the Arctic ice sheet recedes allowing ever-greater access to open water in that region, long dormant maritime boundary disputes between Russia, Denmark, Norway, Canada, and the United States flare up with prospects of access to previously unreachable fossil fuel deposits beneath the continental shelf. Russia's recently adopted National Security Strategy notes the world's increasing rivalry over access to energy resources and does not rule out the possibility that a future competition for energy reserves might require Russia to use military force.⁶³

Escalation Dynamics in a Confrontation between NATO and Russia

Although none of the foregoing issues are likely to result in war, should they or other friction points fester and erupt in a crisis, serious escalatory pressures could arise. Given the geographical proximity to Russia and relative remoteness from the West at which the confrontation would likely occur, NATO's most timely means of response would be to rush air power into the region in an effort to deter or defeat an attack on whichever state Russian forces are threatening. This development would alarm Russian

extended military action ranging from the bombing of the upper Kodori Valley to the deployment of armoured units to reach extensive parts of Georgia, to the setting up of military positions in and nearby major Georgian towns as well as to control major highways, and to the deployment of navy units on the Black Sea." According to the Mission, "All this cannot be regarded as even remotely commensurate with the threat to Russian peacekeepers in South Ossetia." IIFFMCG Report, Vol. I, 2009, pp. 23-24.

⁶¹ Olikier *et al.*, *Russian Foreign Policy*, *op. cit.*, p. 102.

⁶² The latter has been expressed in numerous ways, particularly in Estonia, where in 2007, the removal of a Soviet war memorial from a Tallinn city square and the relocation of graves at the memorial to a military cemetery outside of the city triggered a cyber attack that paralyzed that country's political and financial institutions for several days. The attack was launched from Internet protocol addresses inside Russia, but Moscow denies any official involvement. See Ian Traynor, "Russia Accused of Unleashing Cyberwar to Disable Estonia", *The Guardian*, May 17, 2007.

⁶³ *National Security Strategy of the Russian Federation to 2020*, May 12, 2009, paragraphs 11 and 12.

leaders, as they have observed the United States and its allies unleash high-intensity air campaigns on several adversaries during the last two decades, dismembering air defense systems and striking command-and-control centers with devastating effectiveness. Writing in professional journals, senior Russian military officers have expressed serious concerns about whether Russia could defend itself against such attacks.⁶⁴

Western leaders should consider how Moscow might react to such a threat. One would hope that the potential consequences of escalation would make both sides very cautious, but that is far from certain. During the Cold War, mutual risks of catastrophe helped stabilize several East-West confrontations in Berlin and Cuba. But even in that era the benign outcomes of those events owed almost as much to luck as to effective crisis management. Management of the Cuban missile crisis in particular was a delicate balancing act that could easily have toppled into war had it lasted a day or two longer, or had even one more accident occurred like those on October 27, 1962.⁶⁵

Today the dynamics of an East-West confrontation would be very different. Whereas U.S. atomic superiority was balanced by Soviet conventional superiority early in the Cold War,⁶⁶ and rough nuclear parity

⁶⁴ See for instance V. V. Barvinenko, "A Retrospective Probe into the Growing Role of Armed Confrontation in the Aerospace Sphere", *Military Thought*, Vol. 15, No. 3, 2006, pp. 12, 22; B. F. Cheltsov, "Matters of Air and Space Defense in Russia's Military Doctrine", *Military Thought*, Vol. 16, No. 2, 2007, pp. 51-52; and V. A. Subbotin and A. M. Shavelkin, "Tendencies in the Development of Combined-Arms Combat Tactics", *Military Thought*, Vol. 17, No. 2, 2008, p. 54.

⁶⁵ On that day, at the height of the crisis, a U-2 reconnaissance plane flying in the Bearing Straight strayed over the Soviet landmass and was pursued by Soviet fighters. U.S. fighters scrambled from bases in Alaska to defend it. According to Laurence Chang and Peter Kornbluh, Alaskan Air Command records suggest that the U.S. fighters might have been armed with nuclear air-to-air missiles. Fortunately, the U-2 escaped and all aircraft returned to their bases without further incident; however, aircraft flying reconnaissance over Cuba that same day were not so lucky. One conducting a low-level mission was damaged by Cuban antiaircraft fire. It managed to return to base, but a U-2 was destroyed over Cuba later that day by a Soviet SAM, and its pilot was killed. This latter incident raised pressures from U.S. military leaders for President Kennedy to authorize retaliatory strikes against the SAM sites in Cuba. See Chang and Kornbluh, *The Cuban Missile Crisis*, *op. cit.*, pp. 376-377.

⁶⁶ Western students of Cold War history remember the era of massive retaliation as one in which U.S. nuclear superiority deterred Soviet conventional aggression in Europe, often forgetting the fact that Soviet conventional superiority also deterred the United States from attacking. In both the Truman and Eisenhower administrations, some military and civilian officials argued for launching a preventive war against the Soviet Union before it could build an atomic arsenal large enough to threaten the West. At least two studies were conducted to determine the potential outcomes of such an attack. In both cases, it was determined that even a mass atomic attack on Russia "would not seriously impair the Soviet Army's ability to advance quickly into Western Europe, the Middle East, and Asia." For more on that era's preventive war debate, see Karl P. Mueller, Jasen J. Castillo, Forrest E. Morgan, Negeen Pegahi and Brian Rosen, *Striking First. Preemptive and Preventive Attack in U.S. National Security Policy*, Santa Monica, RAND Corporation, 2006, pp. 121-152.

imposed the stability of mutual assured destruction later on, the United States and its Western allies now enjoy unrivaled conventional superiority.⁶⁷ At first consideration, one might conclude that the approximate parity in today's U.S. and Russian strategic nuclear arsenals would impose stability on any future confrontation similar to that seen in the Cold War, and agreements typified by the April 2010 New Strategic Arms Reduction Treaty (New START) tend to encourage that conclusion.⁶⁸ But whereas strategic parity made Cold War leaders cautious, it could lead today's leaders to place more faith in stability than is warranted. Moscow might gamble on hopes that fears of escalation would make the West shy of confronting a Russian military intervention in a nearby state; Western leaders, in turn, might assume that Moscow's fear of escalation would grant NATO the freedom to use its conventional superiority to defeat such an intervention and promptly impose a military solution on Russia, just as Western coalitions have done on other post-Cold War opponents.⁶⁹ The result of these dynamics could be that both sides find themselves in a rapidly escalating war culminating in a dangerous game of brinkmanship... or worse.⁷⁰

Given a misplaced faith in strategic stability, escalation dynamics in a conflict between NATO and Russia would not hinge on the risks of a strategic nuclear exchange, at least not initially; rather, they would build from the bottom up. NATO intervention in a conflict on Russia's border would likely begin with an air campaign against Russian military targets. This would be very dangerous. During the Cold War the main line at which NATO and Warsaw Pact forces faced off passed through the middle of Germany; neither superpower succeeded in pushing substantial forces onto the other's doorstep.⁷¹ Today, however, NATO members and aspirants shoulder the Russian border, leaving Moscow no protective buffer and precious little reaction time should NATO forward-deploy in a crisis. These

⁶⁷ For a thoughtful discussion of Russian fears of U.S. conventional superiority, see Dennis M. Gormley, "The Path to Deep Nuclear Reductions: Dealing with American Conventional Superiority", *Proliferation Papers*, No. 29, Fall 2009, pp. 27-38, available at: <http://www.ifri.org/downloads/pp29gormley1.pdf>.

⁶⁸ Formally called, "Measures for the Further Reduction and Limitation of Strategic Offensive Arms," the treaty was signed on April 8, 2010, and went into effect on February 5, 2011. The text of the treaty is available at: <http://www.state.gov/t/avc/newstart/c44126.htm>, accessed: June 20, 2011.

⁶⁹ Glenn Snyder first identified this dynamic and called it the *stability-instability paradox*. It posits that if both adversaries possess a reliable and survivable nuclear deterrent, the strategic balance will be sufficiently stable that both of them will be more emboldened to engage in conventional aggression. See Glenn H. Snyder, *Deterrence and Defense*, Princeton, Princeton University Press, 1961 and Glenn H. Snyder, "The Balance of Power and the Balance of Terror", in Paul Seabury, (ed.), *Balance of Power*, San Francisco, Chandler, 1965.

⁷⁰ For a thoughtful discussion on the dangers of misplaced faith in strategic stability, see David S. Yost, "Strategic Stability in the Cold War. Lessons for Continuing Challenges", *Proliferation Papers*, No. 36, Winter 2011, available at: <http://www.ifri.org/downloads/pp36yost.pdf>.

⁷¹ The Cuban missile crisis was Khrushchev's attempt to alter the balance of power by stationing nuclear forces just off the U.S. coast. Resolution of the crisis required President Kennedy's secret agreement to remove the comparable U.S. nuclear threat on the Soviet Union's doorstep, the Jupiter missiles in Turkey.

factors would generate considerable risks of inadvertent escalation, even in a war in which both sides were trying to contain it. Were NATO air commanders to employ their forces in the manner prescribed in U.S. Air Force and Joint doctrine, they would attack Russia's integrated air defense system (IADS) aggressively to establish air supremacy. They would also strike key command-and-control nodes coordinating Russian forces. They would almost certainly carry out such attacks against Russian forces in whatever friendly territory Russia had seized in its initial advance, and given the range and lethality of the Russian IADS, NATO commanders would likely press for permission to strike targets on Russian soil adjacent to the area of operations as well.⁷²

All of this could happen very quickly – indeed, according to U.S. doctrine, the campaign would incorporate elements of “parallel attack” to impose shock and paralysis – which would raise alarm, perhaps even panic, among Russia's military and political leaders. Would Russia respond with a strategic nuclear strike? Of course not. But New START does not place limits on numbers of tactical nuclear weapons, and Russia continues to possess a substantial inventory of them. If sufficiently frightened, Moscow could threaten or even resort to the use of some number of these weapons to offset NATO conventional superiority. Such a response would not be inconsistent with Russia's national military doctrine, which states:

Russia reserves the right to use nuclear weapons in response to the use of force against it and (or) its allies, nuclear and other weapons of mass destruction, *as well as in the case of aggression against Russia with conventional weapons*, which threaten the very existence of the state [my italics].⁷³

In fact, there are indications that Russian military doctrine even envisions using nuclear strikes to “de-escalate” conflicts. In the late 1990s, Russia's long-range aviation element, then the 37th Air Army,⁷⁴ simulated limited nuclear strikes in several of its exercises to show resolve with the expectation that such actions would de-escalate wars with NATO forces. In scenarios simulating responses to large-scale NATO air attacks on Russia, Belarus, and the Kaliningrad Oblast, the 37th Air Army conducted “Tu-95MS and Tu-160 flights over the Atlantic off the Norwegian coast, complete with simulated cruise missile launches against targets in North America and

⁷² See Air Force Doctrine Document (AFDD) 2-1.1, *Counterair Operations*, Maxwell AFB, LeMay Center for Doctrine Development and Education, October 1, 2008, pp. 23-24; AFDD 2-1.2, *Strategic Attack*, Maxwell AFB, LeMay Center for Doctrine Development and Education, June 12, 2007, pp. 9, 30-32; JP-3, *Joint Operations*, Washington, Joint Staff, September 17, 2006, updated March 22, 2010, pp. xxii, III-23, V-13, V-17.

⁷³ *Military Doctrine of Russia*, Moscow, February 5, 2010, Section III-22. An English version of the 2010 Military Doctrine is available on the Nuclear Resonances web page, “Russia's New Military Doctrine in English”, available at: <http://scisec.net/?p=231>, accessed: June 20, 2011.

⁷⁴ In 2010 the 37th Air Army was deactivated. Russia's nuclear long-range aviation capabilities now reside in the Long-Range Aviation command. “Chapter 5: Russia”, *The Military Balance*, Vol. 111, No. 1, 2011, pp. 173-194.

Western Europe”, all, according to Russian news articles, in the belief that such operations, if actually carried out, would de-escalate a conflict with NATO.⁷⁵

So would Russian leaders consider any and all NATO air strikes on its territory a matter that “threatens the very existence of the state”? Probably not, but the answer to that question would depend on the intensity of the strikes, the breadth and nature of targets struck, and how deep in Russia the targets are located. It would also depend on how liberally Russian leaders interpret what threatens the very existence of the state while under the pressure of a high-speed, high-intensity air campaign. Finally, one must consider the question of whether Russia could accept a humiliating military defeat at Western hands, even without attacks on Russian soil, given the importance that Russian leaders place on international prestige and its relationship to state power. Might prospects of such a defeat constitute in their minds a threat to the existence of the state?

If so, Moscow’s resort to the use, or even the threat of use, of tactical nuclear weapons would present a dilemma for Western leaders. The United States would not likely respond to the threat of such an attack with the threat of a strategic nuclear strike, which would risk receiving a strategic strike in return. But neither does the United States maintain a comparable number of tactical nuclear weapons in Europe with which to provide a proportionate deterrent threat.⁷⁶ Facing a Hobbesian choice between risking Russian escalation across the nuclear threshold and acquiescing to Russian conventional aggression in its near abroad, NATO may be forced to accept the latter, which would equate to defeat. Given these concerns, it would be prudent to seriously consider ways to manage the risks of escalation in a conflict between NATO and Russia, regardless of how unlikely such a conflict might appear to be.

⁷⁵ “Russian Heavy Bomber Force Overview”, *NTI*, May 11, 2009, available at: <http://www.nti.org/analysis/articles/heavy-bomber-force-overview/>, accessed: December 30, 2011. Also see, Pavel Podvig, “Russia’s Nuclear Forces: Between Disarmament and Modernization”, *Proliferation Papers*, No. 37, Spring 2011, p. 22, available at: <http://www.ifri.org/downloads/pp37podvig.pdf>.

⁷⁶ Neither the United States nor Russia releases information about the numbers of tactical nuclear weapons it deploys or holds in storage. Estimates in open sources vary, but most put the Russian advantage in tactical nuclear weapons on the order of 10 to 1. For instance the Center for Arms Control and Non-Proliferation estimates the number of U.S. weapons in Europe to be about 200 out of a total inventory of about 500 and the number of Russian weapons deployed to be about 2,000 out of a total inventory of about 5,390. See Center for Arms Control and Non-Proliferation, “U.S. Tactical Nuclear Weapons in Europe Fact sheet”, January 2011, at: http://armscontrolcenter.org/policy/nuclearweapons/articles/US_Tactical_Nuclear_Weapons_Fact_sheet/, accessed: June 20, 2011. Also see Pavel Podvig, “What To Do About Tactical Nuclear Weapons”, February 25, 2010, at: <http://www.thebulletin.org/web-edition/columnists/pavel-podvig/what-to-do-about-tactical-nuclear-weapons>, accessed: June 20, 2011.

Managing Escalation Risks in a Confrontation with Russia

As previously stated, the first step in planning any military intervention should be a serious, objective assessment of the balance of interests between the contesting parties. Such an assessment would, per force, be very contextual, as it would depend heavily on the specific nature of issues in dispute, along with a wide range of historical, political, economic and geographical factors, which together would determine each side's sense of entitlement, vulnerability, and obligation to other interested actors. Such a contextual assessment is difficult to do in any abstract case study; nevertheless, considering the potential routes to conflict illuminated above, one could surmise that a NATO conflict with Russia would likely revolve around the defense of one of the former Soviet republics that are now independent states on Russia's periphery.

Each side's perceived stakes in such a confrontation would depend largely on what states are involved and which side would be trying to change the status quo. In a conflict over one or more of the Baltic states – say, for instance, one triggered by a Russian military incursion into Estonia, ostensibly to punish it for the alleged mistreatment of its ethnic Russian citizens – NATO would likely conclude it has heavy stakes in defeating the Russian attack. Not only are the Baltic states emerging democracies that strongly identify with the West, but NATO credibility would be on the line. Conversely, should some incident between Poland and Belarus flare into a military conflict drawing NATO and Russia into confrontation, Russian leaders would likely feel very invested in Belarus's defense, given that country's cultural and political affinities with Russia and the warm relations and mutual defense commitments between Minsk and Moscow. Which side would perceive the greatest interests at stake in a conflict in Ukraine would depend more on the specific issue at hand and how strongly each side perceives the other to be the aggressor. Perceptions of aggression would also play a role in determining relative stakes in conflicts on Russia's periphery in Central Asia; however, it would be difficult to imagine NATO concluding its interests are great enough to warrant a war with Russia over one or more of those states, given their political, cultural, and geographical distance from the West.

Wherever the conflict might occur, Western leaders should tailor their objectives to match the limited stakes they have in the dispute and resist any temptation to escalate those objectives. The goal should be to blunt Russian military aggression in the global commons or on third-party territory, *not defeat Russia* per se. Military operations should be tailored accordingly. Western leaders should be keenly aware of the escalation pressures that would arise should they conduct operations in ways that threaten the survival of the Russian state or its leaders, or the sovereignty of Russian territory.⁷⁷

⁷⁷ This could be a delicate issue in a conflict involving the Baltic states or Poland, given the proximity of the Kalinigrad Oblast and Russian concerns about its vulnerability.

Once Western leaders understand both sides' stakes in the confrontation at hand, they would need to identify their escalation thresholds. The central premise of threshold management is that escalation can be managed most reliably if every party to a conflict has a clear and realistic understanding of where all important escalation thresholds reside, its own as well as those of other belligerents. To the extent that such an understanding can be achieved, a party can then issue threats and posture forces to deter opponents from violating its critical thresholds while restraining its own actions to avoid transgressing the important thresholds of dangerous adversaries. Conflict then unfolds within the boundaries that each opponent imposes upon the others.⁷⁸

Such an approach in a conflict with Russia would require clear communication at the onset of a crisis, or even earlier. If tensions rise to the point at which it appears that Russia might attack a neighboring state, Western leaders need to decide whether such an event is acceptable. More precisely, they need to decide whether it is so unacceptable that they are willing to pay the cost and bear the risks of military conflict to oppose it. If so, they need to explicitly warn Moscow that its impending action will trigger a Western military response. This warning need not be public – a private communication might be preferable to avoid humiliating Russian leaders – but it does need to be unambiguous. At the same time, Western leaders need to gather as much information as possible in efforts to get a clear understanding of Russian motives and objectives and Moscow's perceptions of each side's capabilities, vulnerabilities, and most importantly escalation thresholds, vis-à-vis the crisis at hand.

The point of issuing an explicit warning to Moscow would be to deter Russian aggression via the threat of conventional military action. To make such a threat credible, NATO would need to posture the forces required to carry it out.⁷⁹ This would be a delicate undertaking. A large mobilization and

⁷⁸ This assumes that one's opponents are indeed dangerous enough that one would want to avoid escalation. In wars in which one side enjoys capabilities largely disproportionate to those of its opponents, as has been the case in some post-Cold War conventional conflicts between Western coalitions and regional adversaries, the party with overwhelming power can sometimes impose its thresholds on its opponents with much less regard for their thresholds. However, this is not always the case if the powerful state's interests in the issue are only marginal and the weaker state's interests are great. Military superiority can only compensate for smaller interests if the stronger party can keep the risks to its forces very limited. Asymmetries of interest tend to result in asymmetries in thresholds. This offers avenues of escalation to the belligerent with the stronger motivation despite its weaker capabilities.

⁷⁹ While nuclear deterrence is achieved via the threat of punishment, conventional deterrence focuses more heavily on threats of denial. In deterrence by denial one postures forces to convince the opponent that one can defeat his attack, or at least make the outcome sufficiently uncertain that it does not appear worth the risks. Glenn Snyder was first to make this distinction in his monograph, *Deterrence by Denial and Punishment*, Princeton, Center of International Studies, 1958. John Mearsheimer and Robert Pape further developed the concept of denial-based coercion, Mearsheimer focusing on its role in conventional deterrence and Pape extending it to compellence. See John J. Mearsheimer, *Conventional Deterrence*, Ithaca, Cornell University Press, 1983, and Pape, *Bombing to Win*, *op. cit.*

deployment of forces would take time; initiating such action might prompt Moscow to attack quickly in hopes of achieving a *fait accompli*, thereby putting the onus of evicting Russian forces on NATO's shoulders.⁸⁰ Moreover, given Russia's fears of Western conventional military superiority, a major effort to rush forces close to the Russian border would alarm Moscow and risk a preemptive strike against those forces before they could be formed for battle. Given NATO's superiority in air power, however, and particularly the U.S. advantage in long-range strike, NATO could quickly deploy air assets to bases around the periphery of the area of operation, sufficiently dispersed to dilute opportunities for Russia to strike them preemptively. Those assets could then impose a threat of air interdiction against any Russian force invading defended territory.⁸¹

Should Russia defy this deterrent and invade anyway, NATO would implement its air interdiction campaign and attempt to blunt the Russian offensive. This would probably impose heavy losses on Russian forces, but it might be costly for NATO as well. Western leaders would want to stop the Russian advance before it penetrated deeply into friendly territory, but engaging Russian forces close to the border would put NATO air assets well within range of Russia's highly lethal homeland IADS. NATO air commanders would almost certainly lobby for permission to strike IADS targets – missile launchers, air bases, radars, and command-and-control centers – on Russian territory adjacent to the area of operations in order to reduce friendly air losses. Whether to allow them to do so would be a tough decision, one that should be based on a careful analysis of how significant a threshold such an escalation would cross. At the very least, one would expect NATO forces to execute non-kinetic attacks on Russian based systems, via electronic and cyber warfare. If those actions degrade the Russian IADS sufficiently to keep air losses down to acceptable levels, then no kinetic attacks on Russian territory should be permitted.⁸²

⁸⁰ A Russian move of this type would be consistent with Schelling's strategy of posturing oneself to put the "last clear chance" of avoiding mutual disaster on the opponent's shoulders, thereby forcing him to back down. See Schelling, *The Strategy of Conflict*, *op. cit.*, pp. 44-45.

⁸¹ This assumes, of course, that NATO air forces could acquire the regional basing and over-flight permissions needed to conduct operations. Such an assumption is reasonable – particularly regarding NATO member states on Russia's periphery – given that NATO would be extending security guarantees to those states; however, it is not assured.

⁸² This, of course, raises questions regarding how to respond to Russian non-kinetic attacks and whether escalation in the cyber domain can be deterred. Electronic warfare (i.e., jamming) has always been difficult to deter, due to the deniability and reversible nature of the attacks. Opinions vary on whether cyber attacks can be deterred, but an analyst who has given the question serious thought doubts that cyber deterrence is viable due to the low costs of entry into this field – anyone with a laptop computer and access to the Internet can be a combatant – and difficulties in attribution – attackers can claim to be civilian "patriots" or remain anonymous. On the other hand, state combatants might be "self-deterred" from using sophisticated attacks on some of the enemy's more important systems at lower levels of conflict due to their ability to penetrate and exploit those systems for the collection of valuable intelligence. A cyber attack *might* create substantial effects against operational forces – the nature and extent of those effects are as

In any event, NATO forces would have to be restrained from crossing the thresholds that Moscow considers most critical. Determining where those thresholds lie would be an ongoing challenge, as Russian leaders might conceive new thresholds as the tide of battle shifts and unexpected vulnerabilities emerge. Western leaders would need to monitor Russian statements and actions closely in efforts to detect subtle changes in threshold perception and adjust their actions accordingly. In fact, communication between the belligerents would be important throughout the conflict. Western leaders would need to persistently warn Moscow through word and deed of where important NATO thresholds lie. Perhaps even more importantly, they would need to assure their Russian counterparts that they intend to respect Russia's critical thresholds: NATO would only be conducting a limited operation to blunt Russian aggression. NATO would have no intentions of taking Russian territory, threatening the safety of Russian leaders or non-combatants, disabling Russia's nuclear deterrent forces, or (and especially) threatening the continued viability of the Russian state.⁸³

Finally, NATO military leaders would need to make concerted efforts to minimize risks of accidental escalation. Straightforward, coherent, and easily understood ROE would need to be developed based on the analysis of Russia's escalation thresholds (and other political considerations), and they would need to be updated as ongoing analyses indicate those thresholds are changing. Compliance with the ROE would have to be closely monitored and strictly enforced. No tolerance for violation is acceptable when at war with a nuclear-armed adversary. Conversely, should incidents of aberrant Russian behavior be discovered, they would have to be carefully evaluated to determine the motives and whether the acts were sanctioned by Russian leaders. Isolated incidents might be considered the crimes of rogue individuals. Alternatively, trends in bad behavior might suggest sanctioned escalation, poorly developed ROE, or a general breakdown in military discipline. In any event, NATO military leaders should not let such forms of escalation go unanswered. If atrocities are committed, they should be publically exposed to impose a political cost on Moscow for allowing them occur. If other forms of apparently accidental escalation occur, NATO leaders should consider measured reprisals to try and bring the conflict back into bounds and deter further transgressions. In

yet unproven – but such effects would probably be, at best, only temporary. At the same time, however, the attack would alert the victim to the vulnerabilities in its system, which it would then quickly eliminate, thereby closing a route of access to the attacker for further attacks or exploitation. Therefore, leaders considering initiating cyber attacks would have to weigh the questionable benefits to be gained in those attacks against the loss of intelligence that would result from them. For more on this discussion, see Martin C. Libicki, *Cyber Deterrence and Cyber War*, Santa Monica, RAND Corporation, 2008; and Martin C. Libicki, *Conquest in Cyberspace: National Security and Information Warfare*, Cambridge, Cambridge University Press, 2007.

⁸³ A critical point in Schelling's work is too often overlooked or forgotten: for a coercive threat to have any meaning, it must be accompanied by the *assurance* that if the opponent complies with it, you will not punish him anyway. "Take one more step and I'll shoot," can be a deterrent threat only if accompanied by the implicit assurance, "And if you stop, I won't." Schelling, *The Strategy of Conflict*, *op. cit.*, p. 74.

doing this, however, Western leaders need to communicate to Moscow the reasons for the reprisals and, if appropriate, its assurance that NATO does not intend to escalate further.

Assuming NATO and Russia could be successful in keeping a conflict within boundaries acceptable to both of them, one side would eventually begin to prevail over the other. This would present an extraordinary challenge for threshold management. As Carl von Clausewitz observed, the drive to struggle ever harder to achieve victory or avoid defeat is a natural human tendency.⁸⁴ Should Russia (or NATO) find itself sliding towards defeat, it would be sorely tempted to escalate beyond the established boundaries in efforts to reverse that trend. At this point, the prevailing side would have to exhibit a combination of firmness and empathy. Assuming NATO is that party, Western leaders would need to sternly warn Moscow that any escalation to the use of nuclear weapons would be met with a catastrophic response, one that would inflict costs on Russia that far exceeded any benefits that could be expected in the struggle at hand. At the same time, however, Western leaders should also express a willingness to settle the war in a way that would not humiliate Russia and, if possible, that would satisfy at least some of the grievances Moscow harbored that led to the conflict. This would require true compromise at some level, which is more than mere face-saving.

Just as the Cold War strategists ultimately realized, if war between nuclear adversaries is to be survived, it cannot be a “winner-takes-all” affair. It must be a limited war, fought for limited gains. At the end of the day, the victor must allow the defeated to depart the field with dignity and limited losses.⁸⁵

⁸⁴ Clausewitz, *On War*, *op. cit.*, p. 77.

⁸⁵ This sentiment was expressed most eloquently, perhaps, by Fred Iklé. See Iklé, “Can Nuclear Deterrence Last Out the Century?”, *op. cit.*, pp. 84-105.

Conclusion

This essay has examined the escalation dynamics that could arise in a military confrontation between NATO and the Russian Federation, and it has described how threshold management concepts might be applied to manage the risks that those dynamics could create. The exercise was conceived as a modest attempt to cover an area not addressed in a previous RAND study⁸⁶. That work explored the escalation risks that could emerge in a U.S. conflict with another large nuclear state, China, as well as with two emerging regional powers, North Korea and Iran, and with non-state adversaries. The findings of this examination regarding Russia largely parallel those of the RAND study's work on confrontations with other state adversaries.

Like Russia, the three states examined in the RAND study are all intimidated by Western, or, more specifically, U.S. conventional military prowess and are looking for ways to offset it. China, unlike Russia and the other two states, is focusing the greater share of its efforts on developing conventional means to challenge U.S. military access to the Western Pacific, such as conventional ballistic missiles targeting U.S. regional bases and, eventually, carrier task forces.⁸⁷ The Chinese are also attempting to develop capabilities to interdict the U.S. space systems that provide force enhancement support to U.S. expeditionary forces.⁸⁸ All of these elements present prospects for rapid escalation, both in the intensity and scope of conflict.

China's greater emphasis on conventional preparations stems partly from its economic capacity to do so, as compared with the other three states, and partly from the fact that Beijing has, historically, maintained a declaratory policy of "no first use" (NFU) regarding nuclear weapons. Despite these considerations, however, Beijing has signaled through unofficial statements by senior People's Liberation Army (PLA) officers that it might "waive" its NFU restraints in certain circumstances, such as if U.S. conventional air power attacks in ways that threaten China's nuclear deterrent forces.⁸⁹ And similar to Russia, Chinese military writings suggest

⁸⁶ This study is reported in the 2008 monograph by Morgan *et al.*, *Dangerous Thresholds*, *op. cit.*

⁸⁷ *Ibid*, pp. 68-71.

⁸⁸ *Ibidem*, pp. 71-76.

⁸⁹ *Ibidem*, pp. 67-68.

that PLA operational doctrines might embrace the concept of using nuclear weapons “for the purpose of deescalation and, perhaps, war termination.”⁹⁰

Other ways that dynamics in a conflict between China and the U.S. might differ from those in one between Russia and NATO stem from differences in stakes. Whereas Russia’s wounded pride and historic sense of entitlement to regional dominance might cause it to stumble into a conflict with the West (or vice versa), China is motivated by a much keener sense of lost sovereignty regarding Taiwan and a deep-seated sense of victimization by the West. To date, Beijing has been relatively patient regarding the island’s unresolved political status and has not made a concerted effort to force the issue. That could change, however, with the shifting correlation of forces, especially if Taipei crosses a Chinese “redline” by overtly declaring independence. Should conflict erupt, Chinese leaders would be highly motivated to see it through to an acceptable conclusion. China watchers have speculated that the Chinese Communist Party might not be able to sustain a humiliating defeat at U.S. hands and stay in power. Even more so than with Russia, Beijing could see prospects of a serious military defeat as a threat to regime survival, and that could be source of rapid escalation.

A confrontation with North Korea or Iran would likely exhibit some of the same dynamics as one with Russia or China, but could be much more volatile. Both of those states are much weaker than China or Russia and are therefore much more afraid of U.S. conventional capabilities. Unlike Russia and China, their nuclear capabilities would be very immature. They would probably lack well-established policies, procedures, and training for nuclear surety and release, their second-strike survivability would be in question, and their leaders would not be experienced in nuclear stewardship or nuclear diplomacy. The RAND study found that new nuclear states have historically tended to overestimate the coercive leverage obtained by nuclear weapons. As a result, several have provoked dangerous confrontations with rival states.⁹¹ The sum of these findings is that North Korea and Iran would be more prone to provoke a confrontation with the United States than would Russia or China and also be more unpredictable in their behavior once they found themselves in a crisis.

All of this suggests that effective threshold management will be crucially important in an armed conflict with any of the aforementioned states. Western leaders will need to assess the balance of interests and identify each side’s critical thresholds. They will need to illuminate these thresholds to opponents in ways that deter deliberate escalation and reduce the risks of inadvertent escalation. They will need to manage their forces firmly to avoid escalatory accidents, and they will need to calmly evaluate and respond to the accidents that will inevitably occur over the course of the war. Most of all, they will need to restrain their objectives and settle for limited gains, which will most likely amount to defeating the opponent’s aggression in ways that simply preserve the status quo.

⁹⁰ *Ibidem*, pp. 61-62.

⁹¹ *Ibidem*, pp. 85-116.

Thankfully, the world has never witnessed a major conventional war between nuclear-armed adversaries, much less one in which nuclear weapons were exchanged. Studies late in the Cold War raised serious doubts whether the latter could be kept limited, or even prosecuted in a coherent manner, given the massive disruptions in communications and physical, mental, and emotional dislocations that would occur at multiple levels of command once nuclear weapons began detonating on each side. Although a handful of analysts continued to lobby for counterforce, nuclear war-fighting strategies to the very end of the era, the ranks of those who accepted Kahn's thesis that nuclear wars could be fought and won had by then grown exceedingly thin in the West and were substantially diminished in the East. The near consensus was that any nuclear war would likely be uncontrollable, resulting in consequences so tragic that victory, however defined, would be pyrrhic.

The implication of such a conclusion is that for any escalation management framework to be viable, it must inform strategy making while the conflict is well below the nuclear threshold. Further, it must face up to the uncertainties inherent in war – the lack of perfect information and perfect control; the subjectivity of perception; the inevitable miscalculations that result from incompetence, fear, and fatigue; and the general unpredictability of human behavior – and offer realistic approaches for managing these factors to the extent they are manageable. Cold War-era approaches to escalation management failed to meet those criteria. As a result, decision makers on both sides of the East-West divide abandoned them and relied instead on conflict avoidance.

Threshold management is a framework for reducing escalation risks in an era in which Western leaders may not be able to avoid confronting nuclear-armed adversaries without abandoning important interests. It seeks to reduce the ambiguity of war by illuminating thresholds and keeping the conflict within bounds of those deemed most critical via strategies fashioned on a judicious balance of deterrence, assurance, and self-restraint. In doing so, it seeks to benefit from the strengths of previous approaches while avoiding their most serious weaknesses. It acknowledges uncertainty and engages in coercive negotiation with the opponent much in the way that Schelling advocated, but avoids irreversible commitments and making threats that leave something to chance. In seeking to keep the conflict bounded at the conventional level of war, where Western military forces have a clear operational advantage, it attempts to achieve escalation dominance much in the way that Kahn envisioned, but it cautions planners and leaders that multiple dimensions of escalation may be available to savvy opponents. It urges its users to be ever vigilant, collecting and analyzing information on emerging thresholds and managing excursions from them. Ultimately, however, it cautions leaders that opportunities to impose escalation dominance will probably be rare. War, after all, is a struggle between human beings, none of whom ever submit to dominance easily.

In attempting to find this golden mean between escalation dominance and brinkmanship, threshold management suffers from some of

the weaknesses of both, and it creates additional dilemmas for strategists. For instance, avoiding irreversible commitments in a confrontation with a nuclear-armed adversary gives decision makers the freedom to change course to avoid calamity. But in doing so, it may suggest to the opponent that one might not be fully committed to carrying out deterrent threats. This is a particular problem when threatening a nuclear response and the opponent has greater stakes in the issue at hand. How can leaders make their threats sufficiently credible without taking steps to make the commitments irreversible? At the other end of the problem, how can Western leaders convince contemporary opponents that they do not intend to decapitate, disarm, or topple their regimes, given the overwhelming military superiority that Western forces now enjoy and will likely bring to bear in a conflict? In essence, how does one make assurance credible while defeating the adversary's military forces? Finally, how does one fight a limited war to its culmination, deterring the enemy from escalating in the face of defeat, or, alternatively, resisting what would likely be enormous pressure to escalate to preclude one's own defeat?

Such questions will loom large in any effort to operationalize threshold management concepts in strategy making, and they offer fruitful avenues for further research.

Information

All published issues of the Proliferation Papers series can be downloaded from the Ifri website.

www.ifri.org

The latest contributions include:

- Bruno Tertrais, “In Defense of Deterrence. The Relevance, Morality and Cost-Effectiveness of Nuclear Weapons”, *Proliferation Papers*, No. 39, Fall 2011.
<http://www.ifri.org/downloads/pp39tertrais.pdf>
- Keith A. Hansen, “Intelligence and Nuclear Proliferation: Lessons Learned”, *Proliferation Papers*, No. 38, Summer 2011.
<http://www.ifri.org/downloads/pp38hansen.pdf>
- Pavel Podvig, “Russia’s Nuclear Forces: Between Disarmament and Modernization”, *Proliferation Papers*, No. 37, Spring 2011.
<http://www.ifri.org/downloads/pp37podvig.pdf>
- David S. Yost, “Strategic Stability in the Cold War: Lessons for Continuing Challenges”, *Proliferation Papers*, No. 36, Winter 2011
<http://www.ifri.org/downloads/pp36yost.pdf>
- Charles D. Ferguson, “Potential Strategic Consequences of the Nuclear Energy Revival”, *Proliferation Papers*, No. 35, Summer 2010.
<http://www.ifri.org/downloads/pp35ferguson.pdf>
- Jing-dong Yuan, “Chinese Perceptions of the Utility of Nuclear Weapons: Prospects and Potential Problems in Disarmament”, *Proliferation Papers*, No. 34, Spring 2010.
<http://www.ifri.org/downloads/pp34yuan.pdf>
- Jonathan D. Pollack, “North Korea’s Nuclear Weapons Development: Implications for Future Policy”, *Proliferation Papers*, No. 33, Spring 2010.
<http://www.ifri.org/downloads/pp33pollack.pdf>

For further information on the *Proliferation Papers* collection, please feel free to contact Ifri’s Security Studies Center: strategie@ifri.org.