The Primacy of Alliance:
Deterrence and European Security

In collaboration with the Atomic Energy Commission (CEA)

Lawrence Freedman

March-April 2013
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.*
March-April 2013

The Primacy of Alliance: Deterrence and European Security

Lawrence Freedman
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.

*N.B. This issue is exceptionally edited and published in UK English.*

**Editorial board**

Editor: Etienne de Durand  
Deputy Director: Corentin Brustlein  
Principal Scientific Adviser: Jean Klein  
Layout Assistant: Fanny Alarcon

**How to cite this publication**

Sir Lawrence Freedman has been Professor of War Studies at King's College London since 1982. He became head of the School of Social Science and Public Policy at King's in 2000 and was appointed Vice-Principal in 2003. He was educated at Whitley Bay Grammar School and the Universities of Manchester, York and Oxford. Before joining King's he held research appointments at Nuffield College Oxford, IISS and the Royal Institute of International Affairs. He was appointed Official Historian of the Falklands Campaign in 1997 and was appointed in June 2009 to serve as a member of the official inquiry into Britain and the 2003 Iraq War.

Contents

Introduction ........................................................................ 7

The Nature and Diversity of Deterrence ......................... 9

    The Three Levels of Deterrence ................................. 9

    From Theory to Practice ......................................... 11

NATO and Nuclear Salience after the Cold War ............... 15

Enduring Issues of Credibility for NATO’s Deterrent posture 19

    Historical Dilemmas of Extended Deterrence ......... 19

    Deterrence and the Primacy of Alliance .................. 21

    A Slowly Eroding Credibility .................................. 24

Conclusion ..................................................................... 27
The “Deterrence and Defence Posture Review”, adopted by NATO at its May 2012 Summit in Chicago, reaffirmed that none “should doubt NATO’s resolve if the security of any of its members were to be threatened”. To this end it would maintain “the full range of capabilities necessary to deter and defend against any threat to the safety and security of our populations, wherever it should arise”. This included nuclear weapons – “a core component of NATO’s overall capabilities for deterrence and defence alongside conventional and missile defence forces”. Happily, the review had “shown that the Alliance’s nuclear force posture currently meets the criteria for an effective deterrence and defence posture”. Perhaps this was because as the review noted, also happily, that NATO did not “consider any country to be its adversary” and that the “circumstances in which any use of nuclear weapons might have to be contemplated are extremely remote”. The alliance did not actually link its nuclear arsenal to those of others. This might be implied by this statement: “As long as nuclear weapons exist, NATO will remain a nuclear alliance”. But it could also simply mean that as long as the United States and, to a lesser extent, the United Kingdom and France, has nuclear weapons they would serve the alliance, irrespective of what others might be doing. Their benefit was presented as being beyond just deterring nuclear use. They provided the “supreme guarantee of the security of the Allies”.1

NATO language has to be decoded. Ostensibly such acts of communication speak to the outside world, proclaiming the alliance’s cohesion and resolve. In practice the main audience is internal. The defence and foreign ministries of member countries check such statements largely to ensure that radical departures in policy have been avoided and that no commitments have been made to anything that will cause upset at home or, even worse, require extra funds. Only rarely do these statements provide evidence of strategic innovation, except for the cognoscenti, who are so sensitive to the nuances of doctrine that they can detect movement that escapes the untrained eye. The cumulative effect, which is intended, is to encourage a sense of continuity as well as shared purpose. There is no desire to use such reviews and communiqués to open up a debate about nuclear weapons and deterrence. Memories of such debates are often painful. They tend to raise awkward questions to which there is often no good answer and may even nudge governments away from alliance orthodoxy.

The language has changed since the end of the Cold War but only cautiously and without ostentation. While the security environment faced by NATO may be quite different from anything either experienced or even imagined during its first four decades, it has been important to assert that its security policies are following the same principles. As the security environment is relatively benign, at least as far as the risk of major war is concerned, this may not matter too much. But suppose that a real crisis developed, would the established policies cope? With not much happening it is tempting to assume that the absence of crisis confirms the wisdom of NATO’s deterrence posture, but surely this cannot be because potential enemies are taking any notice? The language and many of the actual capabilities have been in place for a long time. The words now sound more formulaic and measured than robust. The numbers of weapons are down and attitudes towards nuclear weapons have changed. Can such statements, lacking urgency, direction and any real target, have anything to do with deterrence?

I argue in this paper that such statements do perform a deterrent role because they help hold the alliance together, and it is NATO rather than nuclear weapons that is most essential to deterrence in Europe. Because of alliance with the United States, European members of NATO enjoy a security guarantee that includes protection against a nuclear threat. But this security guarantee strains credibility. This was the case even when few doubted the will of western democracies to use weapons of mass destruction in extreme circumstances. The strain is even more evident now that civilian targets have become far less legitimate in western military thought, and when American foreign policy is increasingly focused on the Asia-Pacific region. If a searching debate on the meaning of a nuclear alliance were to take place, it would throw up a series of awkward questions, and popular engagement might demonstrate that public opinion does not naturally embrace NATO orthodoxy. Perhaps because of this, the inclination in the alliance is to sustain this orthodoxy through regular repetition of standard formulations. Maintaining alliance cohesion may justify bland statements of doctrine but it also creates an awkward paradox. Without any particular challenges regular re-affirmations of alliance unity and devotion to deterrence carry no cost, but they also reduce the incentives to take any steps that might be required to bolster deterrence and sustain unity should a serious challenge materialise. Should it be necessary to deter a real and present danger, would threats carry conviction? If these threats raised the possibility of nuclear war, would the alliance buckle under the pressure? To address these questions I consider the nature of deterrence, the particular problems of nuclear deterrence within NATO, and the impact of changes in both attitudes to the use of force and in geopolitical configurations.
The Nature and Diversity of Deterrence

As a strategy, deterrence has a negative aim and can take several forms, depending on what needs to be protected and the source of the danger. As a deterrent, nuclear weapons provide the ultimate guarantee against threats to national security. The most challenging questions of nuclear strategy revolve around their ability to provide the same degree of security for allies. This is what has come to be expected of NATO since it was founded in 1949, although there are obvious reasons to doubt the credibility of American promises to extend its deterrent “umbrella” in this way. Whenever the future of nuclear deterrence in Europe is addressed this question of credibility is soon thrown into sharp relief.

The Three Levels of Deterrence

Deterrence has worked for NATO because it is a natural posture for a status quo power. The risks fall on any state that wishes to disrupt the status quo because they would be setting in motion a series of events that could end in catastrophe. So long as nothing is happening then deterrence might be said to be working. But there can also be many reasons why nothing is happening, including indifference and inertia as well as restraint induced by explicit, coercive, conditional threats. Nothing happening raises profound issues of causation. Perhaps nothing happens because nobody intends to do anything, or because they lack the ability even if they have the intention. Perhaps there is both intention and capability but something still persuades them that action is a bad idea. That something might be the deliberate threat intended to deter the act, or some part of this threat, or else it might relate to other elements of the situation that are found to be threatening even if not deliberately set up to be so. Thus deterrence can occur on a spectrum. At one extreme a would-be aggressor assesses the risks as too great without the potential victim doing anything, at the other an aggressor holds back because the potential victim has issued a deliberate and compelling warning.

As a foreign policy choice, designed to stop things from happening, deterrence assumes another’s hostility and seeks to deal with it through a coercive threat. This is in preference to alternatives, such as pre-emptive action, support of a third party, preparations for resistance, inducements or appeasement, or just hoping for the best, although it could be combined with some of these. It requires clarity about the vital interests that would lead to the threat being implemented. The offender has to know about a red line which, if crossed, will result in the threat being implemented. Exactly
where this line has been placed may be vague but that should only add to
the caution. The main requirement is a clear relationship between the
prospective offence and the threatened consequences. If that is not
understood then something has gone wrong with the communication of the
threat, which given what we know about human cognition is always
possible. If either side is surprised by a turn of events – the would-be
deterrer at being attacked or the should-be deterred at the ferocity of the
response – then the policy has failed. The threat has not been made, or if
made not noticed, or if noticed not believed.

This relationship between what is to be deterred and how this might
be achieved in international conflict can be simply represented by three
levels – light, regular and plus – which reflect the interests being defended.

At the lightest level the requirement is to prevent action that does
not impinge on the most vital interests. If deterrence of this sort fails it
should not lead to a major war but a more limited action, and for that
reason the deterrent threat is unlikely to involve nuclear weapons. Indeed
because it would be hard to justify the more severe responses when the
stakes are limited, deterrence may fail here because the aggressor views
the risks attached to some probing offensive also being limited.

Regular deterrence concerns the most vital, existential interests, the
safety of the state and a way of life, and for that reason it is often
instinctively linked to nuclear weapons. The deterrent threat posed in this
case would involve major war. It is assumed that a state that could defend
itself with nuclear weapons would do so, especially if the enemy was also
nuclear-armed.

Deterrence plus refers to a situation where an existential threat can
only be deterred by an ally being ready to risk major war. The most extreme
version, which is the one the US adopted during the cold war, threatens
nuclear retaliation to deter a conventional attack on a third party. It is
because of this that NATO had a credibility problem. It required making
threats on behalf of interests that might be deemed less than vital
(something other than defence of the homeland) which if implemented
carried high risks as a result of a counter-threat (retaliation in kind or even
worse). This came to be known as “extended” deterrence.

The impact of a deliberate threat depends on it being taken
seriously by the target. Its credibility could make the difference at a time of
crisis, when relations are tense and an attack may even seem imminent.
Such situations preoccupy the literature on deterrence. Once the crisis has
been managed, and assuming the threat has not been withdrawn, then the
issue of credibility subsides. Patrick Morgan has provided a helpful
distinction between immediate and general deterrence.2 Immediate
deterrence occurs when there is a high risk of hostile action, an explicit and
convincing warning that this would meet a robust response, and then

---

Vol. 41, No. 1, February 1979, pp. 245-247.
nothing happens. If deterrence has been successful and the guard is maintained there should be no incentives for an opponent to provoke a subsequent test. During the early stages of the Cold War there was immediacy to deterrence in Europe because there was deep anxiety about Soviet motives and capabilities. As crises came and went there was greater confidence that they could be managed, especially after the two great confrontations of the early 1960s – over Berlin and Cuba. The superpowers appeared to be moving close to the brink, but their concern about the consequences of war was evident in the readiness to explore concessions and find opportunities to negotiate.

So while deterrence continued to be discussed with apparent urgency, eventually this talk lost its bite. Political leaders did not start each day checking to see if the potential enemy had lowered his guard. Over time the need to avoid major war became accepted by both sides and this was reflected in their behaviour towards each other. I have called this “internalized deterrence”, which occurs when a particular action is considered so imprudent that it has come to be taken for granted and passed into subconscious thought. Any lingering desire to change the status quo is overwhelmed by the recognition that it is too risky to deserve serious consideration. When deterrence has been truly internalized the more provocative courses of action in a given situation are ruled out almost intuitively.

From Theory to Practice

When the strategic value of nuclear weapons was first considered deterrence was not the issue. The original purpose in August 1945 was to end the Pacific War by demonstrating to Tokyo the consequences of a failure to surrender. The demonstrative effect was soon also linked to the embryonic Cold War. Some went so far as to claim that Hiroshima and Nagasaki were not destroyed to end the Pacific War but to strengthen the American hand in a developing struggle with the Soviet Union over the future of Europe. This again presumed a coercive role for atomic bombs, warning Moscow against overplaying its apparently strong hand in establishing a hegemonic position. In practice it was the more traditional forms of power in play at the time that shaped the new configuration in Europe.

In August 1949, as the United States was still building up its stockpile of atomic bombs, the Soviet Union tested its own device. Thus began the arms race which soon led to the development of the even more powerful thermonuclear weapons, followed by intercontinental ballistic missiles (ICBMs). These developments reinforced those who had assumed

---

from the start that the prime strategic value of these weapons lay in their deterrent role. There was talk of pre-emption in the early 1950s, acting swiftly before the Soviet Union caught up, but there was no stomach for such a desperate act and soon the moment when it could have made any sense at all passed. The arms race was animated by the possibility of acquiring a first strike capability, one that would make it possible to disarm the enemy in a surprise attack, or at least fear that the other side might do so. By the mid-1960s the talk was more of mutual assured destruction.

The emerging nuclear stalemate might have been expected to encourage American policy-makers to reduce reliance on nuclear threats to keep the peace, to stay with “deterrence regular”. Indeed the Truman Administration did take this view in 1950, encouraging the members of the new North Atlantic Treaty Organization to develop a conventional capability to match that of the East. The Eisenhower Administration judged this to be both costly and unnecessary. Instead it increased reliance on the nuclear threat, supposedly only until the Soviet Union caught up in practice. In January 1954 Secretary of State John Foster Dulles warned Moscow that any expansionism anywhere could result in “massive”, that is nuclear, retaliation. At the same time West Germany was joining NATO, having agreed to rearm in return for an alliance commitment to protect its Eastern border against further communist encroachments. This had been decided when the objective was to achieve a balance in conventional forces. The Eisenhower Administration’s conviction that it could rely on nuclear superiority allowed it to ease back on conventional forces. Eisenhower was convinced that the rearmament launched a few years earlier was financially foolish; most European countries were struggling to find the resources given the competing demands of post-war reconstruction and the occasional colonial war. As a result the forward defence of West Germany now depended on nuclear deterrence. This dependence became ingrained because no attempt was then made to develop a conventional deterrent, even as the nuclear balance became more equal, and even as the Soviet Union was assumed (erroneously) to be taking the lead in deploying ICBMs. In this way “deterrence plus” found its way to the heart of NATO doctrine.

Many western strategists would have preferred not to be deterring a Warsaw Pact conventional invasion by means of a nuclear threat. A distinction was developed between deterrence by denial or by punishment. Denial meant preventing gain, by means of a demonstrable ability to stop aggression in its tracks, reducing the potential aggressor’s confidence in a successful invasion. Denial came to be associated with a traditional conventional defence. Without denial there was little choice but to threaten a severe punishment. However unattractive this stance was in principle,

---

5 The most important early presentation of the thesis was in Bernard Brodie (ed.), *The Absolute Weapon*, New York, Harcourt, 1946.
7 Glenn Snyder was originally responsible for the distinction in *Deterrence by Denial and Punishment*, Princeton, Center of International Studies, 1958.
NATO countries preferred it in practice, because it was less expensive, and because war in the middle of Europe would be dire even if non-nuclear. All war had to be deterred not just non-nuclear, and a stress on conventional balances reduced the risks for a Soviet Union if it was tempted to try its luck. So given the presumed conventional superiority of the Warsaw Pact, or at least the reluctance to attempt to match it, NATO had to threaten to use nuclear weapons first.

Issues surrounding punishment therefore forged the conceptual framework associated with nuclear deterrence. This framework became progressively more elaborate and complicated as ways were sought to inject some credibility into a first-use threat even though the likely consequence would be devastating retaliation. Generations of strategists were kept busy developing ingenious explanations as to how deterrence could be made to work despite the apparent flaw at its heart and despite the deep antagonism between the two ideological blocs. The explanations always came back to a simple observation: what deterred was not so much specific threats as a general fear of the terrible unknown of another major war. All parties recoiled from a war that would be so horrific that even the notional victors would have to contemplate the ruination of everything they held dear. Nuclear weapons reduced the appetite for war and so provided a natural support for the status quo. “Safety”, as Winston Churchill predicted, had become “the sturdy child of terror and the twin brother of annihilation”.

There were anxieties from the start of the nuclear age that mutual deterrence would break down at some point with disastrous results. Yet it outlasted the Cold War. Instead of a catastrophic Third World War by the 1980s the talk was of a “long peace.” It cannot be proven that this was because of a fear of nuclear war. After two catastrophic wars in a few decades any possibility of a sequel, whatever the types of armaments available, would have generated caution among political leaders on both sides of the Iron Curtain. In an alternative, non-nuclear history, the conventional military balance, considered on its own, might have encouraged once again the thoughts of a swift decisive victory that had prompted past aggressors to risk war. In our actual history, in the few times that political leaders faced a real possibility of nuclear war they appeared terrified. This influenced their behaviour during moments of crisis.

---

8 Winston Churchill, Speech in House of Commons, 1 March 1955, available at: http://www.winstonchurchill.org/learn/speeches/speeches-of-winston-churchill/102-never-despair. He concluded, in what some have seen as his farewell address to the British people: “All deterrents will improve and gain authority during the next ten years. By that time, the deterrent may well reach its acme and reap its final reward. The day may dawn when fair play, love for one's fellow-men, respect for justice and freedom, will enable tormented generations to march forth serene and triumphant from the hideous epoch in which we have to dwell. Meanwhile, never flinch, never weary, never despair.”


To the extent that first use threats had credibility it was not because a way was found round the threat of retaliation but because of the expectation that any conflict that moved further into the dark and intense unknown of a superpower confrontation would carry an increased risk of mishap and miscalculation. The possibility of nuclear escalation would increase as a result of the chaos and irrationality of large-scale conflict. To help the escalatory process, short-range, so-called “tactical”, nuclear weapons were in a position where that could get caught up in a hard fight in the middle of Europe. The threshold into the nuclear apocalypse could be passed as a “tactical” response to the exigencies of battle.
The ideological division of Europe ended over twenty years ago and one of the Cold War alliances has collapsed. Yet though the context has been transformed, and it is now hard to imagine the sort of desperate situation that might trigger escalation to total war, NATO’s core nuclear doctrine and posture have not changed. Deterrence has not been abandoned and so must be presumed to be playing at least a prospective role in European security. The low risk of major war in Europe is still assumed to have something to do with the existence of nuclear weapons, which creates an effect that is still described as “deterrent”. Thus Britain describes its nuclear strike force as “the deterrent”, even though this begs the question of whether deterrence is what it actually achieves and what actually is being deterred.\footnote{For example the White Paper on the future of the UK Trident programme: The Future of the United Kingdom’s Nuclear Deterrent, Presented to Parliament by The Secretary of State for Defence And The Secretary of State for Foreign and Commonwealth Affairs, Cm 6994, London, December 2006.} Although after 1990 there was some vicious internal strife within Europe in the vacuum left by communism’s retreat, major war shows no sign of returning. One reasonable explanation is that the long peace continues because there is nothing left to fight about and so it does not depend on the “balance of terror”. The issue of balance is no longer relevant. The Soviet Union has slimmed down to Russia and has lost the Warsaw Pact: NATO enjoys conventional predominance. Consequently, the main nuclear issue is now spoken of as being more one of disarmament than deterrence.

With conventional superiority NATO no longer needs a first-use threat. Instead it is Russia which now assumes that without nuclear threats it could not cope with NATO’s (or at least the US’s) conventional might. This explains why Moscow is no longer wedded to the “no-first-use” pledge made in the early 1980s when the Warsaw Pact was at full conventional strength. Second, there are now very few short-range American nuclear weapons left in Europe – up to 200 B-61 bombs for dual-capable aircraft based in Belgium, Germany, Italy, and the Netherlands – that somehow are considered to signal American commitment to NATO. This seems small beer compared to the 7,000 such weapons, in all shapes and sizes that were available during the Cold War, as well as the 300,000 American troops. There are current plans to update them. The B61-12 will be a precision-guided air-launched nuclear weapon, carrying a warhead which could be ranging in yield from 0.3 to 45 kilotons. The price tag of probably
$10 billion may make the programme prohibitive, especially if there is lack of enthusiasm in Europe. Current plans are for the German Tornado fleet to be mothballed in 2020, while the successor Eurofighter is not currently designed to carry nuclear weapons. Modification costs would have to be paid by Germany itself.12

The other major initiative is ballistic missile defence. NATO wishes this to be as non-provocative as possible to Russia and geared only to Iran. The system is based on US early-warning satellites and a Turkish-based radar, and a variety of land and sea-based systems, including the new SM-3 interceptor, that could have some capability against limited incoming missiles. The issue has been complicated because of Russia’s insistence that this is a potential threat to its own nuclear deterrent. If it was then that would have significant consequences for the NATO deterrent – in principle making it more credible but also raising the risk of an arms race of sorts. At the moment Russian claims exaggerate the actual and also the potential capability of the new system.

There was an argument during the Cold War that there were some targets that could only be destroyed using nuclear weapons. The lethality of conventional munitions has now reached a point where only very few specific targets could withstand attack. Moreover, claims that alternative forms of mass destruction, such as biological or chemical weapons, could only be deterred by nuclear weapons do not withstand scrutiny although the formal NATO position is that this is a possibility. Conventional weapons can be used in a variety of ways, and while the current presumption is that it is best to avoid populated areas, if it was really desired to retaliate against civilians then that could also be done, although more likely (and appropriately) would be attacks against other things valued by the offender. In the case of Iraq in 1991, for example, although the then Secretary of Defense Richard Cheney refused to rule out nuclear use in response to chemical weapons attacks (unlike the British and the French), Saddam Hussein was warned that the most likely response was that his regime would be toppled. The point is not that this particular threat might be replicated with others but that with attacks lower down the destructive scale there is no requirement to respond in kind. Indeed it would often be foolish and unnecessary to do so, and governments should have a variety of possible responses in their repertoire.

If these last two points are correct, then the assumption that one way or another any future hostilities between nuclear-armed states would escalate to nuclear use, and that escalation would take over as one mega-explosion led to another, may no longer be valid. In an existential war, where the very survival of the state is already at risk, it might be expected

---

that moments of such intense emotion could be produced that even apparently rational political leaders would find themselves giving the launch orders. At the end of the Second World War, after the holocaust, carpet-bombing and V-missiles, the atom bombs seemed to be a logical culmination of what had gone before, and also brutally successful in bringing a total war to an end. The simplest if depressing assumption was that war had become progressively more murderous, with ever more sophisticated means being found to slaughter people on a large scale. There was no reason to suppose that future wars would not be even more intense and existential.

The trend in conventional war since 1945, at least in the West, has been to seek more ethical strategies that deliberately avoid civilians, and refrain from the sort of raids against centres of population that both sides employed during the Second World War, and in later campaigns such as Korea. In part this is because of revulsion at the consequences of city-bombing; in part it is because of a view that even at its height the strategic effects were limited as societies absorbed punishment in preference to surrendering; in part because over time targeting has reached levels of precision unimaginable in the past. We have reached the point where the expectation is that only the intended target should be hit and any collateral damage is unacceptable. This could change. Perhaps under the strain of war attitudes could switch, as they have switched before, into a position where the old arguments about getting at governments through their miserable populations will appear credible again, or there is just a desire for retribution. For the moment, however, it is hard to imagine leaders getting back into a mind-set where nuclear threats could be made with conviction in circumstances other than as retribution for the prior destruction of their countries.

Capping these other factors is the trend in political discourse in the West to delegitimize nuclear weapons. A number of senior figures have come out in favour of “global zero” and similar schemes to abolish all nuclear weapons. This has now been endorsed by President Obama, but as a long-term goal. Political leaders continue to claim that they view nuclear war only with abject horror and often imply that the only justification for their nuclear arsenals is that they can deter the nuclear arsenals of others. They would gladly contemplate complete disarmament, they insist, so long as they could be assured that it was equally complete for potential enemies.

When mutual disarmament was at its most practical, during the early stages of the Cold War when arsenals were still small, prevailing levels of distrust and the fear of cheating ensured that there was no meeting on minds. The situation got more complicated as more nuclear powers appeared on the scene. Grand schemes for nuclear abolition have become steadily more problematic as they require an artfully

14 Remarks by President Barack Obama, Hradcany Square, Prague, Czech Republic, 5 April 2009.
choreographed process of mutually reinforcing and ultimately simultaneous disarmament. By necessity this would involve finding solutions to some of the world’s most persistent and pressing conflicts. More prosaically these proposals can also be realistically seen by countries other than the United States as being entirely self-serving. This is not always appreciated by the main proponents because, to be fair to them, they were pushing similar schemes when they did not so self-evidently serve the interests of the world’s strongest military power. The ease with which the abolitionist movement can now get mainstream support in the United States reflects a view that abolition would not only spare the world the prospect of the instantaneous end of civilization as we know it (not a trivial matter) but also give the United States greater freedom of manoeuvre in its foreign policy. This is why the second most important nuclear power, Russia, has no interest in complete disarmament. It lacks reliable allies, continues to feel vulnerable and considers its own nuclear arsenal as an essential deterrent to NATO.

So we can assume that we will be living in the nuclear age for some time to come. Nonetheless, the cumulative effect of all these changes has been to marginalise nuclear weapons in Western security policies. Few would want to test the proposition that the firebreak between conventional and nuclear war is now more likely to be effective than was assumed to be the case during the Cold War. It is, however, at least arguable that the presumption of escalation from an initial skirmish between major powers to devastating nuclear exchanges that once underpinned the practice of deterrence, and was always speculative, is now even less well-founded than before. This continuing marginalization of nuclear weapons in Europe raises important credibility issues for NATO’s deterrent posture, exacerbated by ongoing political, therefore deeper changes at the heart of the Alliance.
Does this decreasing salience of nuclear weapons within the Alliance threaten the integrity of NATO’s deterrent posture? The problem of credibility the Alliance faced in the past was the result of deterrence plus. The issue was not so much whether Moscow would be deterred so long as it faced the threat of nuclear war but whether the Americans would continue to pose the threat if they were likely to suffer huge losses as a result of Soviet retaliation. More than force ratios and weapons systems characteristics, political relationships are at the heart of deterrence in Europe. That the Alliance’s deterrence credibility lies primarily with the solidarity between its members and the willingness of the United States to remain committed to European security, however, raises serious concern about its future, as we witness a potential long-term reshuffling of Washington’s foreign policy priorities.

**Historical Dilemmas of Extended Deterrence**

The question of whether New York or Chicago should be exposed for the sake of Berlin or Paris was first raised in the 1950s and never went away. In his confirmation hearings in 1959 Secretary of State Christian Herter observed that: “I cannot conceive of any President involving us in all-out nuclear war unless the facts showed clearly we are in danger of all-out devastation ourselves, or that actual moves have been made towards devastating ourselves”.15 A positive answer insisted that American and European security were inextricably connected, but most European governments recognized that whatever might be said for the sake of reassuring allies at times of relative stability, if matters ever came to a head then there could be no guarantees. British Defence Secretary Denis Healey captured the dilemma in what was described as his “theorem”: “it takes only five percent credibility of American retaliation to deter the Russians but ninety-five per cent credibility to reassure the Europeans”.16

If the Europeans had known about the private doubts at the highest American levels about the commitment to use nuclear weapons come what may on Europe’s behalf they would have hardly been reassured. Recently declassified documents from the Nixon era provide fascinating glimpse into

discussions well away from the antiseptic language of NATO communiqués. Consider a discussion in 1969 around a new anti-ballistic missile system. The President observed that: “Suppose you could defend cities. Really means credible threat of first strike would be much greater if they are screwing with Allies”. After Pentagon officials explained that this would still not be available, even with ABMs, the Chairman of the Joint Chiefs of Staff opined: “If I thought technically, fiscally feasible to [develop?] ABM defense which gave first strike capability, I would advocate it, destabilizing or not. Wouldn’t bother me.” To which Nixon responded: “Wouldn’t bother me either. Nuclear umbrella in NATO is a lot of crap. Don’t have it”.17

So long as these things were not said in public and Moscow remained cautious then life was simpler for the Europeans if they declared themselves reassured. The consequences of disbelief were too stark. Only France under General de Gaulle was prepared to accept the logic of doubting the Americans, which meant building an independent nuclear arsenal and leaving NATO’s integrated military command. The British might have privately entertained similar doubts but saw no realistic alternative for NATO as a whole to the American nuclear “umbrella”, however leaky it might turn out to be in practice. The British rationale for an independent nuclear force acknowledged the doubts but claimed to be geared to reducing their impact. Under the “second-centre of decision-making” theory, the British insisted that the problem was not that the Americans would actually let Europe down but a Russian perception that they might. Once the British force was added to the American in the NATO mix then Moscow would have to consider two centres of nuclear decision. Even if they got one right they might still get the other wrong.18 It is not worth bothering about the contrived nature of the claim. The point was not so much about deterrence but alliance, a means of signalling that it was possible to reconcile belief in the US commitment to European security and an independent nuclear capability.

The French were no more able to escape this issue. Once a country had its own nuclear capability a question had to be answered about its attitudes towards the security of non-nuclear neighbours. It was either going to be supremely selfish and remain indifferent to their fate as they suffered aggression, or else it was going to have to commit to extend deterrence and so risk nuclear war even before national boundaries had been violated, unless it accepted that every member of NATO could only really be secure with its own nuclear arsenal. The short-range Pluton missile illustrated the problem. It suggested that France might seek to reinforce deterrence by firing nuclear munitions into German (albeit Eastern) territory. Not only did this imply abandoning West Germany, it also suggested a readiness to initiate nuclear war on German soil should Warsaw Pact armies appear to be headed towards France.

No countries followed France in the 1960s away from NATO’s integrated military command and towards their own nuclear force. They saw few advantages in respect of either foreign or defence policy. The status quo suited them, and it also appeared to have acquired a comforting permanence. US investment in nuclear capabilities had resulted in strategic effects that far outweighed anything that could be achieved by similar expenditure on conventional capabilities. This generated an additional attraction for European governments that went against any yearning for greater independence. A security system that depended on nuclear deterrence was not only largely paid for by the United States, but also reduced the need for conventional forces and so spared their defence budgets. Of course nuclear war would be truly awful, but two sustained bouts of conventional war had also been awful. If the risk of the higher level of awfulness prevented a resort to the lesser then that was all to the good. Protest movements, notably in Britain, denounced the arms race and the prospect of universal nuclear death, but for these reasons governments embraced deterrence.

The political judgement was facilitated by the lack of a dynamic towards war. Once the Cold War settled down in the early 1950s, deterrence became relatively straightforward. Political relationships and tensions had been congealed. At least in Europe, if less so elsewhere, the lines were clearly drawn. Any transgression would be evident and provocative. The ambiguous status of Berlin demonstrated what might have happened if more territory had been contested. During the Berlin crisis of 1961 it was possible to watch an American President developing his risk calculus. War might be contemplated to preserve the freedom of West Berlin but not for the sake of a then hypothetical reunion with the East. Note that this crisis stemmed from an internal weakness in the Soviet bloc – the haemorrhaging of people from the communist East Berlin to the liberal capitalist West. Scenarios for war were regularly developed, often involving surprise attacks, but these lacked verisimilitude because the political triggers were implausible. The most plausible tended to stem from a presumed crisis borne out of instability within the Warsaw Pact.

**Deterrence and the Primacy of Alliance**

The question of the role of nuclear weapons in Europe was therefore always bound up with the question of alliance. It is important to recall that the original act of deterrence intended to persuade the Soviet Union not to expand into Western Europe was not a declaration of nuclear threat but the formation of alliance. The claim leading up to the April 1949 Washington Treaty was that if only the Kaiser in 1914 or Hitler in 1939 had known that they would end up fighting the United States they would have avoided war. The tragedy in both cases was the delay before the Americans accepted that they had no choice but to defend European democracy. The Atlantic Alliance represented a peacetime American commitment to the European balance of power. Thereafter the big questions in Europe were all bound up with the balance of alliance as much as the balance of terror, let alone the balance of conventional military power.
Of the two alliances the Warsaw Pact appeared as the most homogenous, although internal discipline required occasional invasions, while NATO was forever being pronounced as being in “disarray”. The causes of disarray were not just over deterrence. There was for instance the question of differential burden-sharing. Many European states appeared as “free riders”, taking advantage of high American defence expenditures. There were differences of opinion over policies outside the NATO area, notably the Middle East. Yet despite these disagreements the alliance held together because it never served anybody’s interests to pull it apart, once the Gaullist challenge had been effectively contained as France was left isolated. Policies were adopted with the first objective of sustaining an internal consensus rather than maximizing the impact on the external environment. Thus the alliance adopted a position of deterrence and détente,19 combining a hard stance against aggression with a softer readiness to seek ways of easing international tension. The approved strategy was “flexible response”, which confirmed by the heading only that there would be a response to aggression (but no aggressive initiative) and that the response would be flexible. This did not really commit anybody to very much but opened the possibility of a progression through conventional resistance, tactical nuclear use and then on to all-out nuclear war. If all went well none of this need be put to the test, and it had to be assumed that Moscow would take NATO’s posturing seriously, possibly more seriously than it deserved. If alliance was a precondition for deterrence then this priority made perfect sense.

When the crunch came the balance of alliance favoured the West and Soviet communism imploded. The Cold War ended and NATO entered a new stage. As would be expected at times of regional turbulence issues of alliance formation and disintegration came to the fore. This can be dangerous as these are often pre-war activities. Because the balance of alliance is critical in the evaluation of a general balance of power important new combinations can make war more – or less – likely. Whatever the particular triggers, the competitive alliance formation of the first decade of the 20th century created the conditions for the First World War, while the Nazi-Soviet pact paved the way for the Second.

Soviet President Mikhail Gorbachev sought to make the case in 1990 for the dissolution of both Cold War alliances but it was evident that the Warsaw Pact no longer had any residual legitimacy and it soon evaporated. Meanwhile, Germany chose to re-unite within NATO and Moscow acquiesced. The presumption was that NATO would not expand but offer instead looser types of relationship under the “Partnership for Peace” programme. This was not, however, what the Central and Eastern European countries wanted and they pressed for inclusion in NATO. In part this was because of a desire to become an integral part of the West (when accession to the European Union was bound to take longer) but it was also for old-fashioned security reasons. Alliance was seen as a safeguard

---

against future Russian expansionism. Though a number of the western European members of NATO were wary of a move that was bound to be provocative to Russia, the Americans adopted this as policy and so NATO expanded. Russia, without its former fellow Republics of the now dissolved Soviet Union and economically, militarily and politically weak, could do little about it. Nonetheless as the alliance moved towards its borders Russia felt cheated and insecure, a condition from which it has yet to recover. In addition to NATO’s expansionism Moscow became alarmed at its activism, notably the Kosovo campaign in 1999 and then, more ambitiously, the International security Assistance Force (ISAF) in Afghanistan.

We can assume that absent NATO, Russia would have even more turbulent relations with its “near abroad”. It moved into Georgia in 2009, which was then hoping for inclusion in NATO, but was still outside. By contrast Russia’s behaviour was much more cautious with the Baltic states, which were part of NATO. For this reason the newer members of the alliance, close to the Russian border, have worried that NATO’s Afghan involvement has diverted attention from its true purposes. While Russia has no obvious partners if it wishes to form a balancing alliance in Europe, if its foreign policy remains as bad-tempered as it has been over the past decade then a clash with a country in its “near abroad” might lead to the meaning of alliance being explored in an immediate and uncomfortable setting.

Relatively weak states on the periphery of a potentially hostile large power are those most in need of alliance, but for the same reason they also risk dragging those who offer alliance into a war that they might otherwise have avoided. It is one thing to offer defensive protection, quite another to provide cover for provocative, imprudent behaviour. One of the challenges in managing an alliance lies in ensuring restraint among all members. This is an issue that the United States is facing more often in the Asia-Pacific region at the moment than Europe. It was, however, one reason why NATO countries were wary about inviting Georgia into NATO. The most striking example of alliance relationships that generated instability rather than stability when lesser members backing for their bold actions remains the summer crisis of 1914.

European deterrence has always been bound up with questions of alliance and will be so in the future. The key question, therefore, is the durability of NATO. This cannot be taken for granted. Long-established structures can turn out to be very brittle when subjected to sudden stress. It is therefore important to keep in mind the benefits of this unique peace-time alliance. If the Atlantic Alliance did not exist it would now be extremely difficult to invent it. It has been around for so long and is so institutionalized that it may well be that it will be sustained by inertia as well as the lack of any pressing reasons to leave. At the same time, because it may not have many operational tasks in the future, the case for the alliance, and the national military contributions that underpin it, may be difficult to make. This is the basic problem with deterrence. If nothing untoward is going on, which is the objective of the exercise, it may well be assumed that it is a solution to a problem that no longer exists.
A Slowly Eroding Credibility

NATO’s deterrent posture does not feel under strain. It depends on security guarantees first made by the United States to its European partners in 1949 and which have been regularly reaffirmed. We have recalled the public and private doubts about whether these guarantees would withstand a full-blooded crisis but the situation has never arisen. Such a test seems even less likely now. The conviction that another great power war would be such a disaster has been so internalised that it is no longer seen as a serious answer to a great power dispute.

While this security guarantee has been in place the world has changed substantially. There are new types of political enemies and new types of military capabilities in place. Precisely because we do not know exactly what makes deterrence work, can we assume that just because security tensions did not lead to a major war during the Cold War the current set of tensions will also not lead to war? If Europeans found themselves in the midst of an unexpected crisis, might there be a really unpleasant surprise as they discovered that the sense of restraint that had kept the peace in the past was no longer so firmly in place?

It is obviously the case that the prevailing forms of inter-state tension are quite different from those of the Cold War and, by and large, while often vicious and dangerous they do not threaten to escalate to the scale of a great power war. One leftover from the cold war is the partition of Korea. Recent North Korea actions have provided reminders that the 38th Parallel is only an armistice line and that a state of war could resume, this time with a nuclear-armed and even more isolated North. The April 2013 crisis tells us something about the credibility of nuclear threats. No country has ever made such blood-curdling and vicious nuclear threats, promising pre-emption and merciless devastation. Yet to Pyongyang’s consternation they were not taken seriously. This was not because of the small size of the arsenal or the North’s uncertain technical proficiency but because it was extremely hard to see what it could possibly gain from a nuclear war. The assumption was that the threats were part of a desperate attempt to extort from South Korea and the US economic and political concessions. The challenge for the North is that it has no assets other than its nuclear power, but it dare not bargain that away because it would then have nothing with which to bargain or deter its enemies. The other two important points about the crisis were the continuing danger of miscalculation, so that some provocative incident could lead to a more general and unwanted confrontation. The second point was the importance of China, vital to keeping North Korea afloat and clearly unhappy about its reckless behaviour. Washington’s crisis management depended on working in concert with Beijing.

Yet at the same time the rising power of China poses its own challenges to American power. Arguably the most challenging fault lines in the international system are around China’s contested maritime boundaries, as evident in the South China Seas or the spat with Japan over the Senkaku/Diaoyu islands. This is where the United States may have to face up to some significant risks if it is to meet its commitments to allies.
Elsewhere conflicts often reflect intra-state instability. These create uncertain challenges for western states which have a capacity to intervene but are fearful of the long-term and often painful obligations these interventions may entail. Nuclear weapons may be relevant in some of these cases. Iran would still be seen as a radical, challenging factor in the wider Middle East without the prospect of it “going nuclear” but this prospect adds to the sense of danger and urgency. If Iran did become a nuclear power then issues of extended deterrence would be raised in a new context, notably with the Sunni states such as Jordan, Saudi Arabia and Egypt who have appeared particularly concerned about a “Shia bomb”, but also those Gulf states who are now close to the United States, and provide them with facilities. Although Israel has been most vocal on the issue in principle it should be able to cope better because it has its own nuclear arsenal.

These various Middle Eastern challenges continue to exercise a pull on American foreign policy, but the most important policy change under the Obama Administration has been the explicit shift of attention from the Atlantic to the Asia-Pacific region. From an American perspective this makes sense. In economic terms this is a region of growing importance and the American position is coming under challenge from China. At a time when defence budgets are under pressure, priorities had to be made and the choice is not surprising. In a way it is a testament to the success of NATO that the continent is seen to be becalmed, at least in terms of great power conflict. But for Europeans there is an underlying anxiety. Troop numbers are continuing to drop – from 80,000 to 70,000 by 2017, including two brigades now deployed on a rotational basis – and there is a clear US expectation that European countries will accept greater responsibility for their own security and that of their backyard. In his valedictory address as Secretary of Defense Robert Gates told Europeans in no uncertain terms that they must act as serious allies if they wished the United States to take them seriously, but defence spending under austerity continues to fall, and the appetite for operational military roles has reduced commensurately.

The prevailing conditions of austerity do not encourage confidence that defence budgets will hold. The Obama Administration has insisted that it remains committed to NATO but there is undoubtedly a loosening of ties, as the Americans shift their focus and the Europeans become more introverted.

---

20 The White House Office of the Press Secretary, “Remarks By President Obama to the Australian Parliament”, 17 November 2011.
21 “The blunt reality is that there will be dwindling appetite and patience in the U.S. Congress – and in the American body politic writ large – to expend increasingly precious funds on behalf of nations that are apparently unwilling to devote the necessary resources or make the necessary changes to be serious and capable partners in their own defense.” The Security and Defense Agenda (Future of NATO), As Delivered by Secretary of Defense Robert M. Gates, Brussels, Belgium, 10 June 2011, available at: http://www.defense.gov/speeches/speech.aspx?speechid=1581. The valedictory of his successor, Leon Panetta was somewhat gentler. “Remarks by Secretary Panetta”, King’s College London, 18 January 2013, http://www.defense.gov/transcripts/transcript.aspx?transcriptid=5180.
Conclusion

What conclusions can we draw from all of this about the future of deterrence in Europe?

The first is that current security arrangements are likely to endure because there is no better alternative that is also plausible. Questions will, however, increasingly be asked as to whether the alliance is fit for purpose.

Second, so long as it does endure then deterrence is in place, because the combined military weight of the allies could overwhelm all potential challengers, although progressive cuts on the NATO side could erode the differential (assuming that Russia can revive its own armed forces).

Third, this is not to say that all negative events can be deterred. The most problematic will be those that come into the category of deterrence-light, involving ambiguous situations, challenges to less-than-vital interests, and responses that are hard to anticipate in advance and are often tepid or indirect. For example some cyberattacks may come into this category. They are irritating and often damaging but can also raise complicated issues of attribution and impact, and retaliating in kind might not seem appropriate.

Fourth, nuclear deterrence can only confidently be geared to matters of national survival and therefore major war. NATO is in a better position to sustain deterrence than before, not because nuclear threats are inherently more credible, for they are not, but because only a small possibility of nuclear escalation should be sufficient to confirm the inbuilt dread of great power war. There is no dynamic pushing NATO to war with Russia and its conventional superiority means that it is no longer necessary to threaten first use. This is not, incidentally, an argument for a no-first-use pledge, which would not be taken seriously in a big crisis: in the circumstances in which this would become an issue, nobody would dare assume it would be binding. At any rate in the event of such a crisis it would not be helpful if key players concluded that because of this pledge certain courses of provocative action were safe.

Fifth, nonetheless, the problem of extended deterrence remains. Although it may no longer be necessary for the US to threaten a nuclear riposte to conventional aggression it still is expected to threaten retaliation for nuclear aggression. Without the US nuclear umbrella the non-nuclear members of NATO would have no answer to nuclear threats. The only
alternative would be the much smaller British and French forces. The British allocate their nuclear force to NATO but the principles of extended deterrence have never played much role in the UK debate and most of the population would be surprised to find that the Trident force had anything other than a purely national role. So while declining confidence in the overall US commitment to Europe might reinforce the rationale for the national deterrents, London or Paris would be disinclined to talk up a wider alliance role for their nuclear forces to replace those of the United States, and might not be taken that seriously if they tried.

The optimistic but not unreasonable view is that an absence of any serious tendency to major war avoids any serious pressure on the US commitment to Europe, which in turn sustains deterrence and so reduces further the risk of major war. Should this virtuous cycle be broken at any point, as a result of some unexpected crisis, then it could quickly move in the opposite direction. The challenges may at first be ambiguous, a probe to test alliance resolve and cohesion, rather than a full-on assault. Some action in one of the Baltic states may appear inflammatory or a crisis elsewhere in Russia’s near abroad may lead to violence that becomes difficult to contain.

Should relations between NATO and Russia deteriorate further then policy-makers would be far well less equipped to address the fundamental issues of risk and responsibility than were their predecessors. In the years following the Second World War there was a keen understanding of the potential for tragedy inherent in any conflict and concepts of national sovereignty, alliance and security were much less fuzzy than they have become in the interim. Decades of efforts, by their governments and their allies, have gone into shielding the German population from the sort of stark situations that led to such dangerous choices in the past. This population has become increasingly opposed to both civilian and military nuclear power. How would it react to a situation where it faced an antagonistic Russia without the American security umbrella? This is not likely but nor is it inconceivable.

The picture I have sought to develop in this paper is not one of a European security system in crisis. If there is a current crisis in Europe it is because of the stresses and strains caused by the persistent Euro-zone crisis and the associated austerity, or the proximity to instability in the Middle East. My argument is more that a long-established security system that has worked extremely well is slowly being hollowed out. The assumptions about the American commitment and nuclear deterrence upon which the European democracies depended for so long have not been abandoned, and nor are they likely to be, but on close examination they no longer look so robust.

It is nobody’s interest to bring these issues to a head. The history of NATO is one of muddling through. Alliance cohesion has been preferred at each stage to intellectual clarity or even policy consistency. New programmes, especially in the nuclear arena, that risk exposing potential divisions and conflicts of interest are rarely welcomed. This is why it is
unlikely that a successor will be found to the B-61 bombs and why ballistic missile defences are for the moment largely at sea. So long as these are seen as second-order matters then it is not problematic, but the danger is that can encourage within Europe a dangerous complacency. Precisely because there are so many other pressing issues, and because money is so tight, European members of NATO will be content to avoid spending on defence and do no more than acquiesce in nuclear deterrence. It is unrealistic to expect any bold initiatives, but it is not unreasonable to encourage governments to at least remind people about the significance of alliance, the special problems posed by nuclear weapons, and the meaning of security. If they rely on the sort of routine statements with which this paper opened, affirming established security arrangements in bland, bureaucratic language while avoiding all the difficult issues, should a serious crisis erupt they will lack the intellectual framework and language that will enable them to address and explain the challenges they face.
All published issues of the Proliferation Papers series can be downloaded from the Ifri website.

www.ifri.org

The latest contributions include:


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