Topical Essays
The Importance of Alliances for U.S. Security

Martin Murphy

“No man is an island, entire of itself,” wrote the English poet John Donne in 1624. The same is true of nations.

The United States now sits at the apex of an international network of alliances brought together during the Cold War, but this has not always been America’s situation. In earlier times, especially at its inception, the U.S. benefited from alliances, generally as the junior partner. Success in the Revolutionary War was helped by a crucial alliance with France, a country that the infant U.S. shortly thereafter fought in the undeclared Quasi-War (1798–1800).

It is true that George Washington, in his Farewell Address of 1796, warned his countrymen that they should not “entangle our peace and prosperity in the toils of European ambition,” an admonition that has come to be viewed as a warning against “foreign entanglements.” But while he urged Americans to take advantage of their country’s geographical isolation from the world’s troubles, he was not advancing an argument for political isolationism. If anything, he was anticipating (and sharing) the sentiment of British Prime Minister Lord Palmerston, who, speaking in the House of Commons on March 1, 1848, avowed that “We have no eternal allies, and we have no perpetual enemies. Our interests are eternal and perpetual, and those interests it is our duty to follow.”

Washington’s argument, like Palmerston’s, was that no nation, especially a nation as influential at various times as the United States or Great Britain, can disengage from the world. Such a nation must instead be free to choose when to engage and when not to engage—and, most momentously, when to go to war and when to walk away.

Wisdom and Utility of Alliances

An equally spirited debate about the wisdom and utility of alliances continues today. Repeatedly, alliances are referred to as burdens, an elastic term that can be stretched to include everything from moral hazard to free riding.

The burden of moral hazard is that states, including states of roughly equivalent weights, may feel emboldened to pursue riskier foreign policies because their allies are obligated to come to their rescue. Perhaps the most famous example of what is also referred to as “entrapment” was Germany’s alliance with Austria–Hungary before World War I. Emboldened by this alliance and German encouragement, Austria–Hungary felt that it could safely make humiliating demands of Serbia even though Serbia was allied to Russia. It was wrong: Russia failed to restrain Serbia and initiated military preparations of its own, the chain gang of alliance obligations snapped into place, and Europe found itself on the way to war.

The reciprocal of entanglement is abandonment. The U.S., for example, is at risk of
being pulled both ways in its relationship with allies in Asia, a concern that Beijing is evidently attempting to use to its own advantage.\(^8\)

Concerns about free riding, “that America’s allies, especially the smaller ones, have simply been unfair in not bearing large shares of the common burdens,” has bedeviled America’s relations with its allies—especially its NATO allies—for many years.\(^9\) In straightforward economic terms, the U.S. does make a greater contribution to alliance resources than other members, and there is a risk that this could become unsustainable during a period when America’s economic power is in relative decline. However, the costs of alliances, including the sometimes disproportionate cost of alliance leadership, must not be weighed against cash savings but rather against the cost of possible conflict in blood as well as treasure without them.\(^10\)

America’s treaty with France committed it to joining France in war if it was attacked by Great Britain. Since 1792, France had been engaged in its own revolutionary war with its neighbors, including Britain, and the political grouping led by Thomas Jefferson and James Madison was arguing strongly that the United States should fulfill its treaty obligations. Washington, who issued his 1793 Proclamation of Neutrality (subsequently the Neutrality Act of 1794) to avoid this obligation, wrote his address in part to deflect their criticism of his actions.\(^11\)

The Royal Navy was now much stronger than it had been when it was defeated by the French at the Battle of the Virginia Capes in 1781, the action that had precipitated Cornwallis’s surrender at Yorktown, thus ending the War of Independence. Washington, well aware of Britain’s renewed naval strength, refused to see American trade ravaged and U.S. ports set ablaze.

Unlike Madison, who when President launched the War of 1812 that saw the White House burned and, as the naval historian Alfred Thayer Mahan memorably recorded, grass grow in the streets of Boston as a consequence of the British blockade, America’s first President had no intention of exposing his country to such peril.\(^12\) He recognized that the young republic lacked the military wherewithal to deliver on its treaty promise even if it wanted to and assessed that the costs of joining France in a protracted conflict with Great Britain far outweighed any potential benefit for America. The gap in capabilities between the young United States and Britain and the geographic distance separating America from France were simply too great.

The United States and Great Britain concluded no formal military alliance during the 19th century. There were several disagreements, some severe enough on occasion for both sides to contemplate war prior to what historians have called “The Great Rapprochement” between the two beginning in the 1890s,\(^13\) but even before that, there was also complementarity in their actions that accorded with the principle of eternal interests rather than eternal allies. For example, the Monroe Doctrine, set forth by President James Monroe in 1823 to prevent European nations from colonizing territory or threatening states in North or South America, might have been largely impossible to implement given the Royal Navy’s ability to intervene when and where it chose.\(^14\) Britain, however, elected not to challenge the Monroe’s policy because it accorded with Britain’s interest in ensuring that the disintegrating Spanish empire in the Americas did not fall piece by piece into the hands of its imperial rivals.\(^15\)

Clearly, America has chosen to engage in or refuse alliance depending on its interests. So what are the benefits of military alliances if, on occasion and between some powers at least, solemn agreements can be ignored, while in other situations, so much can apparently be achieved in their absence?

**Alliance Typology**

Alliances have been a fact of international political life since antiquity.\(^16\) They perform a number of different functions for states, often at the same time, which makes categorization difficult. Nonetheless, their primary function
is military, and the three primary classifications used in the academic literature bear this out:

- **Defense pacts**, by which signatories are obliged to intervene militarily on the side of any treaty partner that is attacked militarily;

- **Neutrality and non-aggression pacts**, which obligate signatories to remain militarily neutral if any co-signatory is attacked (neutrality pacts are usually more specific than non-aggression pacts); and

- **Ententes**, by which signatories agree to consult with one another and potentially cooperate in a crisis, including one involving an armed attack.\(^{17}\)

The common features shared by all three types of alliances lead to a definition like the one proposed by Stephen Walt: that alliances are formal or informal commitments for security cooperation between two or more states. “Although the precise arrangements embodied in different alliances vary enormously,” Walt argues, “the defining feature of any alliance is a commitment for mutual military support against some external actor(s) in some specified set of circumstances.”\(^{18}\)

Viewed in this loose way, alliances can be either formal, written treaties or informal, unwritten agreements based on anything from tacit understandings to verbal assurances. These, however, may be good enough. Formal agreements have often said little about actual commitment. The Franco–American treaty sidestepped by George Washington, for example, provided more assurance that support would be forthcoming than turned out to be the case. The French sense of betrayal was one of the factors that contributed to the Quasi-War. On the other hand, America’s alliance with Britain before Pearl Harbor was largely tacit, even secret, but nonetheless very real.

Alliances exist to advance their members’ collective interests by combining their capabilities—which can be industrial and financial as well as military—to achieve military and political success. How these are combined can vary, as the academic classifications suggest.

The degrees to which alliances are institutionalized also differ. Most alliances throughout history have been loose, often ad hoc arrangements and subject to the vagaries of fortune and commitment. Most European alliances, such as the various coalitions that Great Britain assembled to defeat Napoleon, were of this type.\(^{19}\) The French emperor was defeated only when the coalition participants finally realized that if they were to free themselves from endless conflict, they had to stand together rather than cut deals for short-term advantage.

Ad hoc alliances often contain strange bedfellows. Britain, a constitutional monarchy with laws passed by Parliament, established common cause with autocratic Russia to defeat Napoleon. Similarly, in World War II, the Anglo–American democracies found it necessary, if they were to defeat Nazi Germany, to join forces with Stalin’s totalitarian state, which had been their enemy and would be again. Throughout the conflict, each side was suspicious that the other might cut a separate deal with the German dictator, and the desire to ensure that neither side did so sustained the alliance as much as military capability did. In fact, as Robert Osgood argues, “next to accretion, the most prominent function of alliances has been to restrain and control allies.”\(^{20}\)

Most alliances are, to some degree at least, asymmetrical. When it comes to commitments, one signatory may expect less of the other militarily. For example, the 1839 Treaty of London in which Britain guaranteed Belgium’s neutrality, while not a military alliance, was necessarily a one-sided commitment by Britain to come to Belgium’s aid if it was invaded, a commitment that Britain honored in 1914.\(^{21}\)

When it comes to capabilities, alliance members can likewise make very different
contributions. Britain's input to the defeat of Napoleon was primarily financial and naval; apart from Arthur Wellesley's campaign in Spain and victory at Waterloo, few British troops were involved. In fact, it was a classic demonstration of how maritime powers achieve their victories.

In World War II, despite the ferocity of the fighting on the Eastern Front and the beaches of Normandy, the war in Europe was won by Anglo-American air and sea power, which crushed Germany's ability to prosecute the war. Arguably, the Red Army would not have prevailed over the Wehrmacht absent the combined bomber offensive and the British convoys that fought to deliver American war matériel to Archangel and Murmansk. Despite Stalin's bombast and demands for a second front, he was probably aware of this truth.

**Cold War Alliances**

When the United States considered how the post–World War II world should be organized, it thought first of collective security institutionalized in the United Nations. This accorded with its core value of democracy and the liberal ideal that international organizations were a way to transcend national differences and antagonisms. However, in geopolitical terms, the U.N. turned out to be a concert of the great powers that sit on its Security Council, each one of which holds a veto over its decisions. With the sole exception of the Korean War, when a U.N. force under U.S. leadership repelled the North's invasion of the South in the absence of a Soviet veto, the United Nations was quickly shown to be an inadequate bulwark against Soviet expansion.

Realizing this, the U.S. sought an alternative way to respond to Soviet adventurism, adopting a policy of containing the Soviet Union politically and militarily. This was enunciated in the 1947 Truman Doctrine and formalized in alliance terms with the foundation of the North Atlantic Treaty Organization (NATO), often referred to simply as “the alliance,” in 1949.

NATO started with relatively modest ambitions that accorded with America's historical antipathy to entanglements. The initial strategy was for an integrated defense of the North Atlantic area in which the Europeans would contribute the land forces while the American contribution would be confined largely to naval force and strategic bombing. However, post-Korea, the alliance rapidly became more complex as the Cold War with the Soviet Union evolved. Maturing into a “highly institutionalized alliance with elaborate decision-making procedures and an extensive supporting bureaucracy” with its own military command structure, it gained the solidity to outlast the defeat of the Soviet Union, its original antagonist, and retain just enough of its military and organizational capability and capacity to oppose that antagonist when it shed its Communist ideology and rediscovered Russian nationalism.

The arguments for NATO's creation were several. Perhaps most important, it made clear that a free Europe was a vital American interest and made manifest America's commitment to Europe's defense. If Europe had been overrun by Soviet forces, this would have compromised two of America's eternal interests: retention of its continental integrity by undermining control of the sea and air approaches to America's eastern seaboard and preventing the Eurasian landmass from being dominated by a single power.

The arguments against NATO arose out of American ideals:

- Alliance membership, and especially the commitment to Article Five, allegedly compromised the nation's freedom of action contrary to the U.S. Constitution in that “an armed attack” against any signatory would “be considered an attack against them all” requiring the provision of all necessary assistance, including the use of armed force.
- It also allegedly undermined the United Nations and the principle of collective
security by accepting the validity of military alliances and what internationalists regarded as the discredited notion of power balancing.30

Between 1948 and 2014, the United States accumulated some 66 defense commitments,31 including commitments to NATO members (the Washington Treaty of 1949) and adherence to a second, multilateral treaty, the Rio Treaty of 1947,32 which took in most countries in Latin America. The U.S. is also linked in formal alliances to South Korea (with which, like NATO, it shares a military command structure) and Japan, Thailand, the Philippines, Australia, Liberia, and some small Pacific island states that previously were U.S. territories.33

In the 1980s, the U.S. created a new category of alliance called “major non-NATO allies” (MNNA), primarily to ease arms transfers and facilitate military cooperation.34 States in this category include Afghanistan, Argentina, Bahrain, Egypt, Israel, Jordan, Kuwait, Morocco, New Zealand, and Pakistan. In 2015, President Barack Obama announced his intention to designate Tunisia an MNNA. Meanwhile, Congress proposed that Georgia, Moldova, and Ukraine should be extended MNNA status following Russia's 2014 invasion of Crimea, and President Obama similarly proposed, following a 2015 meeting with the Gulf Cooperation Council, that the same offer should be made to Oman, Qatar, Saudi Arabia, and the United Arab Emirates, presumably to soften the blow of the upcoming nuclear détente with Iran that was signed later that same year.

While it is conceivable that U.S. protection might be extended to some countries on this list if they were attacked, there is no guarantee that any military measures would be forthcoming. The standing of some is particularly problematic: Pakistan, for example, which is still linked to the U.S. by the 1954 Mutual Defense Assistance Agreement but has moved closer to China (while the U.S. has moved closer to Pakistan's rival, India), and Saudi Arabia, with which the U.S. has close ties but no formal alliance.

The most problematic relationship of all is with Taiwan. U.S. government intentions toward Taiwan have been mired in uncertainty ever since diplomatic recognition was switched from the Republic of China (ROC) to the People's Republic of China (PRC) on January 1, 1979. Even though this ambiguity has persisted through successive Administrations, the U.S. Congress has always maintained a keen interest in the continuation of contacts and preservation of Taiwan's status consistent with the will of its people. The Taiwan Relations Act came into force in 1979 to govern unofficial relations between the two states. Official military relations, however, were essentially ended on January 1, 1980, when the U.S. terminated the U.S.–ROC Mutual Defense Treaty.

Post–Cold War Changes

Two trends characterize the period since the fall of the Soviet Union:

- NATO's enlargement and search for a new raison d'être and
- The preference for “coalitions of the willing.”

The fall of the Berlin Wall in 1989 triggered a wave of popular uprisings that drove Communist regimes from power across Central and Eastern Europe, culminating in the dissolution of the Soviet Union itself in December 1991. Even before the final collapse occurred, NATO's counterpart in the East, the Warsaw Pact, had disbanded itself at a ministerial meeting held in Budapest in February 1991.

Historically, when a threat disappears, the military alliance assembled to confront it folds its tent and leaves. Instead, and almost instinctively, all of NATO's member governments felt that the alliance should continue without, as Stanley Sloan put it, being “fully agreed as to why.”35 Some officials argued that it was more than a military alliance: It was a
community of values transcending any specific military threat. Others were more specific, suggesting that although the Soviet Union was going through its death throes and the Russia that was reemerging appeared to be moving closer to the West, this could change, and Russia could adopt a threatening posture in the future. Finally, and most broadly, NATO was a source of stability. The investment that had been made in physical infrastructure and the pooling of organizational and cooperative experience was too good an insurance policy against future threats to European security to let go.

However, events in the 1990s unsettled alliance relations.

- The first event was NATO’s initial post–Cold War Strategic Concept. Issued in 1991, it emphasized a broader approach to security. In effect, the alliance now needed to manage not one but two core missions: collective defense and “out of area” security tasks ranging from crisis response to military-to-military engagement, which together were more complex militarily and diverse politically than its previously singular Cold War purpose.36

- The second, enlargement of the alliance by the admission of previously Warsaw Pact powers, was a source of contention from the very beginning. While it removed the stain of Yalta, the U.S. was concerned that it would strengthen nationalist factions in Russia that were already suspicious of Western intentions.37 These reservations were to be borne out when Russia invaded Crimea and the Ukraine in 2014. In addition, the populations of Central and Eastern Europe that had direct experience of Communist and Russian rule were adamantly opposed to the idea that Russia was entitled to absorb them into a sphere of influence simply to appease its own historic sense of insecurity and great-power entitlement.

- The third was the wars in Bosnia and Kosovo that gave the world the term “ethnic cleansing” as Croats and particularly Serbs used violence to disaggregate ethnically mixed communities with the aim of creating ethnically homogeneous and contiguous areas. Although both conflicts were precisely the type that NATO’s new strategy was intended to defuse, failures in the alliance’s performance on the ground—particularly its inability to prevent the genocide committed at Srebrenica in 1995—pushed America to implement a bombing campaign that drove the warring factions to sign the Dayton Accords by the year’s end.38

Differences between Europeans and Americans, particularly over the Balkan wars, became so acute that, Lawrence Kaplan suggests, the sides drew as far apart as they had been during the Suez–Hungarian Uprising crises of 1956.39 All that held them together was their representation on the Contact Group, a diplomatic device quite separate from NATO that had been created originally to give a voice to Russia in recognition of its traditional role as Serbia’s ally.40 These divisions effectively paved the way for America’s adoption of so-called coalitions of the willing in the early years of the 21st century.

**Alliance Management**

All great powers that have entered into alliances have encountered problems that have required sometimes enormous diplomatic skills to overcome. An overwhelming external threat often concentrates allied minds, but not always: The British assembled five coalitions against revolutionary France and Napoleon before the sixth defeated him not once but twice. The difference was political maturity. As Richard Hart Sinnreich has written:

> The cohesion of any coalition depends on each participating nation’s self-restraint, above all that of the most powerful... That self-restraint is the more necessary the closer the coalition
comes to achieving its military objectives, when the proximity of victory tempts the stronger power or powers to go it alone rather than accommodate the inconvenient preferences of weaker partners. In repeatedly subordinating the desirable to the attainable without forfeiting the central aim of a Europe free of domination by a single untrammeled will, the authors of the Sixth Coalition revealed statesmanship of a high order. 41

The United States managed its Cold War alliances, for the most part, with great skill, but it was helped by the fact that it faced a great threat:

As long as the Soviet arsenal of nuclear weapons and superior manpower on the ground remained in place NATO’s solidarity was assured... Notwithstanding mutual displays of annoyance, Europeans regarded the American commitment to the Alliance for almost two generations as a guarantee of stability in the West. 42

That sense of overwhelming danger was not strong enough in Asia to prevent the Southeast Asia Treaty Organization (SEATO) from dissolving itself in 1977. It had also dissipated in much of Europe by 1992 when the Balkan Wars broke out, leading to a reawakening of the belief that collective security was preferable to collective defense. For some states, including at that point the United States, Operation Desert Storm in 1991 was a powerful reassertion of the importance of the U.N. and a model for what could be achieved in a world that elevated collective security above narrow state interests. There was even a sense that, potentially, the door was now open for the U.N. Security Council to reassert the military role that the antagonism between the great powers (with one opportunistic exception) had rendered impossible for 45 years.

By 1998, the United States was exploring how, under certain circumstances, the alliance could extend its mandate beyond collective defense in the absence of a U.N. mandate. The 1991 Gulf War, for example, had been mandated by the U.N., but the main players involved in the fighting had been NATO powers, and while the coalition formed specifically for the war was an ad hoc creation, the whole campaign had given the impression of a NATO operation.

These discussions, which took place in the context of a planned revision of NATO’s Strategic Concept, were caught up in the controversy over NATO’s role in the Kosovo War. Although in the end, and in the face of the threatened Russian and Chinese vetoes, the operation went ahead without U.N. approval, France insisted that NATO continue to acknowledge the primacy of the Security Council and, in the European context, the “essential role” of the Organization for Security Cooperation in Europe (OSCE), which had been established to monitor compliance with the 1975 Helsinki Accords. Despite this, the door was left open for the allies to operate without a U.N. mandate in the future. 43 Thus, America’s membership in NATO has given it options to act with partners even in cases where broader consent or support vis-à-vis the U.N. is problematic.

The September 11, 2001, terrorist attacks on the United States triggered a powerful reaction from the international community and among America’s alliance partners.

- The U.N. Security Council passed two separate resolutions condemning terrorism;
- NATO invoked Article Five (an attack on one is an attack on all) for the first time in its history;
- The NATO–Russia Permanent Joint Council condemned the attacks and promised to cooperate;
- Australia invoked the Australia–New Zealand–United States (ANZUS) Pact and instructed Australian personnel to deploy with U.S. forces as necessary;
- The Organization of American States (OAS) invoked the Rio Treaty; and
• Japan departed from post–World War II practice by authorizing its self-defense forces to assist U.S. forces, albeit in a limited number of non-combatant roles.

America's efforts over many years to foster wide-ranging alliances in various forms and with a multitude of partners resulted in an outpouring of support from friends around the world. The U.S. declined most of these offers of support, and this rebuff went down especially poorly with several NATO partners in Europe. The reasons were certainly not straightforward. *The Washington Times* reported that, “according to Undersecretary of Defense Douglas Feith, the United States was so busy developing its war plans in the early stage of the conflict that it did not have time to focus on coordinating Europe’s military role.”44 In the same article, NATO expert Stanley Sloan was quoted as saying that Washington “may have been wrong about the potential utility of at least making a nod in the direction of the NATO offer and using it as a platform for future construction of a more relevant role for the alliance.”45

The real reason may have been that, scared by their experiences working with NATO in the Balkans, U.S. officials were reluctant to be drawn into a ponderous and consensual decision-making process, while the political leadership viewed NATO's offer as a thinly veiled attempt to gain some sort of institutional control over its response to the attacks.46 However, the U.S. did make immediate use of NATO E-3 surveillance planes to monitor American domestic air space and in 2003 gave NATO command of the (by then United Nations-mandated) International Security Assistance Force (ISAF) in Afghanistan.47

**Coalitions of the Willing**

It has always been necessary to measure the cost of alliances against their advantages. By the first decade of the 21st century, the United States appeared to view the costs of formal alliances as too high. The gulf that emerged in the 1990s between America's technological capabilities and those of every one of its allies was in some cases so big as to be unbridgeable. U.S. forces struggled to be able to work with some of them. On top of that, some allies no longer valued a U.S. connection as highly as they once did because the threats they faced appeared to them to be less serious.

To long-standing American complaints of allied free riding—letting the U.S. pay for their defense so that they could spend money on social welfare or economic projects—was added a new complaint: If alliance memberships do not help to ensure that allies do not actively oppose U.S. policy decisions, what are they good for?48 Arguments with European allies over Bosnia and Kosovo, for example, or U.S. withdrawal from the Philippines in the years following the fall of Ferdinand Marcos, or the continuing opposition to the U.S. base footprint on Okinawa all left question marks in American minds about the value of formal alliances.

Alliances are inseparable from their contexts. The world was changing. The context was no longer the Fulda Gap but events in far-off places that, while they concerned the world’s sole surviving superpower, could be of little relevance to other members of the alliance or, for that matter, any static, geographically specific grouping of states. The fear that a spark in some distant brushfire war could ignite a global conflagration had gone. But America could not be so sanguine, and when attention switched to the Middle East, what it needed was not battle tanks but basing rights everywhere from Saudi Arabia to Uzbekistan.49

In November 2002, President George W. Bush announced at a NATO summit that the United States would lead a “coalition of the willing” if Iraqi President Saddam Hussein refused to surrender his weapons of mass destruction (WMD).50 The model was akin to the sheriff’s calling for a posse: It was the mission that decided the coalition, not the coalition that decided the mission. If NATO could not be persuaded to support U.S. foreign policy objectives in Iraq *en bloc*, then individual
members could band together in a coalition whose legitimacy in this case derived from the fact it was made up of free, democratic states. However, that was not essential: All that was required was a common interest or perception of the threat perception and a willingness to do something about it.

Another coalition of the willing but not a military alliance is the Proliferation Security Initiative (PSI), also initiated in 2002. It has now been endorsed by 105 countries interested in preventing the spread of WMD.51

Such coalitions, military or otherwise, are “limited associations of convenience [that leave] countries free to pick and choose specific issues, locations and moments for cooperation based on their individual calculations of the national interest” without requiring them to subscribe to any set of common values or political philosophy.52 They put Realpolitik at the service of America’s predominant liberal internationalism, reinforcing the point that states do not have eternal allies, only eternal interests.

What coalitions of the willing do not do, as Kurt Campbell has pointed out, is institutionalize and encourage habits of cooperation and deep engagement, characteristics that embodied NATO’s operating style during the Cold War and America’s formal alliances like those with Japan and South Korea.53 Relying exclusively on global coalitions of the willing may give the United States maximum flexibility, but it will be in exchange for an increased share of the military burden.54 In Europe and perhaps in Asia, where political and military burdens can and should be shared, it may therefore be premature to call time on alliances, which for nearly three-quarters of a century have been among America’s greatest strategic assets.

**Alliances: America’s Great Strategic Advantage**

Since 1941, “alliances have proven to be a crucial and enduring source of advantage for the United States.”55 How so?

- **Alliances prevent war.** Not every war, of course, but by driving up the cost of aggression, defensive alliances have an effective record of deterring revanchist states from using violence as a means of settling disputes or gambling on a quick military thrust to achieve relatively risk-free advantage. History suggests strongly that states with allies are less at risk of attack than those without them, an observation borne out by the success of U.S. alliances during the Cold War.

This does not mean that aggressors will refrain from using other means to achieve their objectives; in fact, they already are doing so, and campaigns designed deliberately to remain below the level of violent confrontation are likely to become more common. General Valery Gerasimov, chief of the Russian General Staff, has observed that in recent conflicts, non-violent measures occurred at a rate of four to one over military operations and that objectives previously viewed as attainable by direct military action alone could now be achieved by combining organized military violence with a greater emphasis on economic, political, and diplomatic activity.56 Defensive alliances will therefore need to extend the breadth of their activities to avoid being outflanked by opponents that use unconventional means to acquire political advantage.

- **Alliances control rivals.** The United States is first and foremost an air and naval power. It wins its wars by retaining control of its own movement and access to supply and denying similar freedom to its adversary. To do that successfully requires a global network of bases and the ability to control the world’s key chokepoints. Geography and the current U.S. basing structure mean that China, Iran, and Russia are likely to be bottled up in any future conflict—although China’s recent island-building activity in the South China Sea...
reveals a determination to secure its trade routes to the south and west and overcome what has been termed its “Malacca dilemma,” and using non-military means has enabled it to confuse and blunt an effective U.S. and allied response to this expansion.

- **Alliances control allies.** Entrapment is a concern for any dominant alliance partner. Germany failed to restrain Austria–Hungary in 1914—indeed, encouraged it to act quickly to win what it expected would be a short war. This risk makes management of alliance relations essential, something at which the U.S. has proved to be remarkably adept. Conversely, the U.S. has felt constrained on occasion by its alliance partners, but mostly when they were being asked to operate in ways that were removed from the alliance’s primary task.

- **Alliances enable balancing.** When regional states attempt to disrupt the status quo, smaller regional states will either balance against it in an effort to retain their independence or join it (“bandwagon”) in an attempt to curry favor and, by being seen as friends, retain sufficient influence over its actions to limit damage to their own interests. A core of U.S. allies in each region can act as a center of attraction around which balancing can be built, as is occurring now in East Asia. Without them, the sole option for regional powers may be to bandwagon with the regional aggressor.

- **Alliances prevent alliance formation by others.** Most of the world’s military powers are members of U.S. alliances. If these alliances did not exist or were abandoned, states would almost inevitably be drawn closer to China, Russia, and Iran and possibly into alliances in active opposition to the United States.

- **Alliances control the bulk of the world’s military power.** The nations that are allied with the U.S. spend around $1 trillion on defense (about 62 percent of global military expenditure) and have 6 million people (31 percent of their populations) under arms. China, Iran, and Russia collectively spend roughly 17 percent of global defense expenditure and are able to draw upon around 19 percent of global military manpower (roughly 3.7 million people under arms).

- **Alliances can hold the line.** In a multipolar world in which a reduced U.S. defense establishment might have to face multiple threats, strong and confident allies can hold the line even if they may not be able to roll back the aggression by themselves. This allows the U.S. time to prioritize threats and respond when it is able to do so.

- **Alliances facilitate global power projection.** The United States is isolated geographically behind two great oceans. To be able to exert power in Asia, the homeland of revanchist power, it requires bases in Europe, the Middle East, and East Asia. From these bases, it can exert influence and power where and when it needs to do so and in small packets early on to deter and prevent challenges from arising that later could be defeated only by the application of overwhelming force. The notion that the United States could mount a campaign using long-range U.S.-based air power or the concept of prompt global strike alone is based on a misunderstanding of what both capabilities are designed to achieve.

- **Alliances are the cost-effective option.** Preserving peace and sustaining the global political and economic system’s current U.S. orientation can be achieved most cost-effectively with allied support. The alternatives would call for either the maintenance of a huge U.S. military presence overseas far in excess of what is
being maintained now or the holding of substantial forces in readiness at home in case the need arose to fight their way back into Europe or Asia to confront trouble in support of what is called “offshore balancing.”

- **Alliances enhance international legitimacy.** They mean that the United States never has to walk alone. When it resists aggression, it is able to do so with the moral authority of the free world.

**The U.S., Allies, and a Free World**

The free world: a phrase that unfortunately has dropped out of fashion since the end of the Cold War yet is as relevant as ever. China, Iran, and Russia are revanchist powers. All three aim to revise the existing order in their respective regions unilaterally and at the least possible political and military cost to themselves. America is the leader of the free world, and revanchist powers know that if they are to succeed, they must diminish U.S. power globally and undermine the tenets of the current, American-led global order.

Each successful step they take along that path diminishes U.S. security and the security of U.S. partners and allies who accept the current global order as one that serves their own political and economic interests as much as it serves those of the U.S. To achieve their aims, the leaders of China, Iran, and Russia are suppressing individual liberty in their own countries, isolating their populations from information that undermines their control, and concentrating power in their own hands. America has seen the world darken this way before and knows that a darker world is one in which conflict is more likely.

That conflict is arguably underway already: China, Iran, and Russia all act as if it is. In such circumstances, as Winston Churchill put it memorably in 1945, “There is only one thing worse than fighting with allies, and that is fighting without them.”
Endnotes:


11. The U.S. was also engaged in the negotiations that resulted in the Jay Treaty of 1794, which established trade relations with Great Britain, resolved some outstanding military and debt issues left over from the 1783 Treaty of Paris that had formally concluded the Revolutionary War, and is credited with averting potential war between the two countries. This treaty was also opposed by Jefferson. For more information, see George Washington, letter to Thomas Jefferson, April 12, 1793, in *The Writings of George Washington*, Vol. X, Part 4, ed. Jared Sparks (Boston: Russell, Shattuck, and Williams, 1836), p. 336.


16. Two of the most famous examples are perhaps the Delian League, which evolved into the Athenian Empire, and the Peloponnesian League, made up of Sparta and its allies, which contested the Peloponnesian War between 431 and 404 BC. For more information, see “Peloponnesian War,” Livius, last modified March 28, 2016, http://www.livius.org/articles/concept/peloponnesian-war/ (accessed May 17, 2016).


24. Collective security exists when all states pledge to defend the security of all other states under international law. Collective security arrangements are systemic. They involve all—or almost all—states that constitute an international system. Axiomatically, the threats they seek to prevent arise for the most part from within the system, not externally. The United Nations is a collective security system. It and systems like it are not alliances. In contrast, collective defense arrangements are alliances whose members pledge to defend all other members from collectively acknowledged attack arising from outside the alliance. In practice, collective defense arrangements work to defuse disagreements between members in order to maintain internal cohesion.


26. Osgood, Alliances and American Foreign Policy, p. 43.


34. 22 CFR §120.32—Major Non-NATO Ally.


38. Sloan, NATO, the European Union, and the Atlantic Community, pp. 93–97.


40. Ibid., pp. 190–192.


45. Ibid.


49. Jung, “Willing or Waning?”

52. Jung, “Willing or Waning?”
54. Ibid.
The Reality of Cyber Conflict: Warfare in the Modern Age

Paul Rosenzweig

Consider a fairly typical incident from 2014. In March of that year, The New York Times reported a persistent cyber threat, known by the code name “Snake,” that had infiltrated the cyber systems operated by the Ukrainian government. The program gave its operators full remote access to the compromised systems, which allowed the attackers to steal information as well as insert additional malware to create further harm. Citing confidential U.S. government sources, the newspaper attributed Snake to Russian actors and connected the deployment of the Snake virus to Russian intelligence collection and disruption of Ukrainian command-and-control systems.¹

At the same time, of course, Russian troops were on the ground in Crimea, and the potential for kinetic conflict between Ukrainian and Russian military forces loomed. Russia formally annexed Crimea just a few weeks later and since then has rather brazenly supported “separatists” in the Eastern Ukraine.

That single episode captures the new reality of military operations in the cyber domain in many ways. At a minimum, cyber conflict will be part of combined operations against physical opponents. Cyber tools will partake of the character of both espionage activities and traditional military activities. At times, the effect of cyber tools may be equivalent to kinetic weapons; at other times, they will be used in a more limited manner to degrade, disrupt, or destroy data and information. In some cases, the origin and source of the tools used in a cyber conflict will be difficult, if not impossible, to discern, rendering attribution of responsibility for an attack problematic; in others, the origins are likely to be crystal clear but the long-term effects of the tool obscured. And all of this will occur at a time when legal norms about appropriate conduct in cyberspace are in a state of flux, without settled definition.

Perhaps even more confusingly, the nature of the conflict in the cyber domain may diverge from settled patterns of military conflict. We will, of course, likely see conflict between nation-states, but we will also see nation-states in conflict with non-state actors and, oddest of all, can also anticipate conflicts in the cyber domain between two non-state parties. How these conflicts will manifest themselves and the nature of the American military response to them will vary significantly in each context.

State vs. State

In a state-vs.-state conflict, we are likely to see cyber activity coupled with conventional operations. For example, since 2014, the cyber-enabled nature of the Russian–Ukrainian conflict has morphed even further. A partial list of cyber activities associated in open-source media with the conflict between Russia and Ukraine over Crimea and Eastern Ukraine would include:
• Russian pre-attack cyber espionage and network mapping of Ukrainian systems;

• Degradation of Ukrainian telecommunications links to Crimea during the Russian invasion, followed by the severing of cross-border telecommunications connections;

• Russian social network sites blocking sites and pages with pro-Ukrainian messages;

• Russia Today (the Russian English-language website) being hacked with the word “Nazi” prominently inserted into headlines to describe Russian actors;

• An IP-telephonic attack on the mobile phones of Ukrainian parliamentarians;

• Russian forces jamming cell phones, severing Internet connections with Ukraine, and seizing telecommunications facilities in Crimea;

• Multiple hacking operations under the #OpRussia and #OpUkraine hashtags including recruitment operations among local cyber-capable actors;

• A large-scale DDoS attack on Russian websites including the Kremlin and the Russian central bank;

• Similar DDoS attacks on Ukrainian news sites, most noticeably during the Crimean “independence” vote, using the DirtJumper botnet; and

• Noticeable activity by hackers of Turkish, Tunisian, Albanian, and Palestinian origin, more commonly attacking Russian sites in support of Ukraine.

One aspect of the conflict worthy of commentary is the evident restraint by both parties. It appears, for example, that no efforts have been made to have a kinetic, destructive effect on critical infrastructure on either side of the border.

But that does not mean that the critical infrastructure is immune. To the contrary, Russia has been strongly implicated in an attack that took six Ukrainian power companies offline. The power outage was caused by a sophisticated attack using destructive malware known as BlackEnergy, which wrecked computers and wiped out sensitive control systems for parts of the Ukrainian power grid. The attack was so severe that it knocked out internal systems intended to help the power companies restore power. While the power generation systems themselves were not attacked, controlling computers were destroyed, and even the call centers used to report outages were knocked out.²

**State vs. Non-State**

Sometimes a state may be confronted by actions by a non-state actor (or perhaps a putative non-state actor whose activity cannot be convincingly attributed to a nation). Consider the recent late 2014 intrusion at Sony, which provides an instructive case both for testing the limits of our understanding of the legal definition of war and for demonstrating that the laws of armed conflict are not the only means of addressing cyber intrusions.³

The intrusion, conducted by a group identified as the “Guardians of Peace,” exfiltrated terabytes of data from Sony. Some of the data involved unreleased films; other data included embarrassing internal e-mails and proprietary information. Additionally, the hackers demanded that Sony withhold from release *The Interview*, a movie depicting the assassination of North Korean leader Kim Jong-Un. After delaying the release for several days, Sony eventually made the movie available through several alternate outlets. The FBI (relying in part on information provided by the National Security Agency) attributed the intrusion to North Korean government agents.⁴ Sony is not saying how great the damage to its financial interests is, but estimates range upward of $50 million.
Here we have a state actor, North Korea, or its non-state affiliates using cyber means to degrade the economic interests of the citizens of another nation, the U.S. How shall we characterize this action? It had no kinetic effects, nor did it significantly affect the American economy. No matter how we view it, Sony is not “critical infrastructure” of the United States (though, oddly enough, the Department of Homeland Security does characterize it as such), so this is not an “armed attack” triggering the laws of armed conflict. Nor is it even an act of espionage. But calling this a state-sponsored criminal act seems to trivialize its geopolitical context.

In the end, the Sony intrusion and Russia’s disruption of the Ukrainian power grid seem to reflect a new category of conflict: a quasi-instrumental action by a nation-state or its surrogates that has significant but non-kinetic effects on a target nation. Such “attacks” are not a “use of force” or an “armed attack,” but they are likely to generate reciprocal responses from the target state that involve a wide array of state powers. The United States, for example, has publicly announced financial sanctions against North Korea and may very well have taken other, non-public actions in response.

Individual vs. State

Then we have the case of a well-placed or technically proficient individual “attacking” a state, often from inside an organization in much the same way a mole would operate to conduct espionage for a foreign intelligence service. In many ways, this insider threat is the most challenging for a nation because it takes advantage of asymmetric attack capabilities that are especially pronounced in the cyber domain.

Consider the following question: What or who has been the most significant cause of damage to the national security of the U.S. through cyber means in recent years? By any absolute measure, the most likely answer is Edward Snowden—a single individual who, through his own activities or perhaps with a small cadre of a few fellow travelers, caused immense damage to American national security interests. The consequences of Snowden’s actions in 2013 include:

- Major damage to formal diplomatic relations between the U.S. and numerous countries identified as targets of U.S. surveillance or “cyber snooping”;
- Popular outrage among U.S. allies and friends in Europe over what they perceive as egregious American spying against their own national security interests (even though people generally accept that spying occurs even among friends, it becomes a different matter when it is revealed so publicly); and
- Opportunities for countries like China and Russia to create a perception of false equivalence between the nature of what they are doing (rampant economic espionage) and what the United States has been doing (more traditional national security intelligence activities).

Even worse, Snowden disclosed intelligence sources and methods to the detriment of the United States. As a result, terrorist groups and other governments have changed their communication activities so that the U.S. cannot as readily intercept their communications and understand their plans. China, for example, was alerted to a particularly significant penetration of one of their cyber systems—a penetration that, presumably, has since been terminated.

The scope of the damage caused by Snowden is nearly incalculable, and he did it as an independent actor rather than as an agent of a foreign government, which in past times would have been critical to his ability to operate at this level. Advances in the cyber domain have made it possible for individuals or small groups operating unaffiliated with any nation-state to cause profound, national-level damage that would have been unthinkable in
previous eras. And as non-state entities, they have no sovereign interest that might be leveraged as would be the case in a conflict between states.

Therefore, when we look at cyber conflict and threats to national security, we should not focus exclusively on other national opponents. Rather, our cyber conflict strategy needs to account for the “democratization” of conflict in and extending through the cyber domain, by which we mean simply that the tools and weapons of attack are now widely available and that the use of force—and in the context of modern societies, information is very much a tool of force—is no longer the exclusive province of nation-states.

**Non-State vs. Non-State**

In this light, the U.S. is in the midst of what scientist-philosopher Thomas Kuhn would call a paradigm shift. It is a shift that is empowering individuals to act with force in ways that were beyond our conception a few short years ago. To see one example of how that paradigm shift operates in practice, reflect on what we might call the “WikiLeaks War” from 2010—a conflict exclusively between non-state actors—and what role (if any) a national government might have in such a conflict.

With the disclosure of classified information from American sources like Chelsea (née Bradley) Manning, WikiLeaks appeared to be launching an assault on state authority and, more particularly, that of the United States, though other governments were also identified. Interestingly, the most aggressive and decisive response came not from government, but from the institutions of traditional commerce. There is no evidence that any of the governments ordered any actions, but the combination of governmental displeasure and clear public disdain for WikiLeaks Editor-in-Chief Julian Assange soon led a number of major Western corporations (MasterCard, PayPal, and Amazon, to name three) to withhold services from WikiLeaks. Amazon reclaimed rented server space that WikiLeaks had used, and the two financial institutions stopped processing donations made to WikiLeaks.

What followed might well be described as the first cyber battle between non-state actors. Supporters of WikiLeaks, loosely organized in a group under the name Anonymous, began a series of distributed denial-of-service (DDoS) attacks on the websites of the major corporations that they thought had taken an anti-WikiLeaks stand, flooding the websites with “hits” to prevent legitimate access to them. The website of the Swedish prosecuting authority, who is seeking Assange’s extradition to Sweden to face criminal charges, was also hacked.

Some of the coordination for the DDoS attacks was done through social media, such as Facebook or Twitter. Meanwhile, other supporters created hundreds of mirror sites, replicating WikiLeaks content, so that WikiLeaks could not be effectively shut down. The hackers even adopted a military-style nomenclature, dubbing their efforts “Operation Payback.”

When Anonymous attacked, the other side fought back. The major sites used defensive cyber protocols to oppose Anonymous, rendering attacks relatively unsuccessful. The announced attack on Amazon, for example, was abandoned shortly after it began because the assault was ineffective. Perhaps even more tellingly, someone (no group has publicly claimed credit) began an offensive cyber operation against Anonymous itself. Anonymous ran its operations through a website, AnonOps.net, and that website was subject to DDoS counterattacks that took it offline for a number of hours.

In short, a conflict readily recognizable as a battle between competing forces took place in cyberspace, waged almost exclusively between non-state actors.

Anonymous’s failure to target corporate websites effectively and its relative vulnerability to counterattack are likely only temporary circumstances. Anonymous and its opponents will learn from this battle and approach the next one with a greater degree of skill and
a better perspective on how to achieve their ends. Many of their more recent attacks, such as the effort to shut down the Vatican’s website, have already shown a great deal more sophistication and effectiveness.

Moreover, Anonymous has demonstrated that even with its limited capacity, it can inflict significant damage on individuals and companies. When Aaron Barr, corporate head of the security firm HB Gary, announced that his firm was investigating the identity of Anonymous participants, Anonymous retaliated by hacking the HB Gary network (itself a significantly embarrassing development for a cybersecurity company) and taking possession of internal e-mails that suggested that HB Gary was engaged in some questionable business practices. As a result, Barr was forced to resign his post.

More to the point, Anonymous has made quite clear that it intends to continue to prosecute its cyber war against the United States, among others. “It’s a guerrilla cyberwar—that’s what I call it,” says Barrett Brown, 29, a self-described senior strategist and “propagandist” for Anonymous. “It’s sort of an unconventional asymmetrical act of warfare that we’re involved in, and we didn’t necessarily start it. I mean, this fire has been burning.”

Or consider the manifesto posted by Anonymous, declaring cyberspace independence from world governments: “I declare the global social space we are building together to be naturally independent of the tyrannies and injustices you seek to impose on us. You have no moral right to rule us nor do you possess any real methods of enforcement we have true reason to fear.” In February 2012, Anonymous went further by formally declaring “war” against the United States and calling on its citizens to rise and revolt.

In many ways, Anonymous conducts itself much as an opposing military organization might conduct itself. In February 2012, for example, it was disclosed that Anonymous had hacked into a telephone conversation between the FBI and Scotland Yard, the subject of which was the development of a prosecution case against Anonymous. That sort of tactic—intercepting the enemy’s communications—is exactly the type of tactic any government or insurgent force might use, and by disclosing the capability, Anonymous successfully created uncertainty about how much else it might be intercepting.

In advancing their agenda, the members of Anonymous look somewhat like the anarchists who led movements in the late 19th and early 20th centuries, albeit anarchists with a vastly greater network and far more ability to advance their nihilistic agenda through individual action. And like the anarchists of old, they have their own internal disputes, thus making comprehensive or singular analysis of objectives, methods, and potential points of leverage quite difficult. In 2011, for example, another group called Black Hat effectively declared war on Anonymous because it disagreed with the Anonymous agenda.

Even more important, however, Anonymous and its imitators look like the non-state insurgencies that the U.S. has faced in Iraq and Afghanistan: small groups of non-state actors using asymmetric means of warfare to destabilize and disrupt existing political authority.

A Strategy for Cyber Warfare

What are the implications of this paradigm shift for cyber/military strategy? They appear to be profound.

From Russia and China, we can expect some form of rationality in action. We can understand their motivations. We know why the Chinese are stealing intellectual properties to jumpstart their economy. We can make some judgments about what would and would not annoy them. In the end, they are rational actors just as the Russians were during the Cold War.

In the cyber domain, by contrast, the motivations of the actors are as diverse as the number of people who are there, and the closer you look, the more unclear things become. There are indeed many actors with many different motivations. They are often characterized as irrational chaotic actors. Perhaps it is a little
unfair to call them chaotic, but what seems to unify them is disrespect for authority, for hierarchy, for structure, a dislike of it and an effort to work outside of it. In this structure, they look much more like insurgents than national military forces.

Given this evolving shift from primary state actors to the n-player world of cyber warfare, a compelling case can be made for a new strategy that is relevant to the changed security environment. There are three factors that should guide thinking about a new cyber strategy—factors that are remarkably similar to those that shape counterinsurgency strategies.

• Cyber warfare favors asymmetries. Non-state actors with power nearly equal to the power of governmental actors are going to be the rule, not the exception. They can serve as proxies for state actors, as the Russian “patriotic hackers” do, but they are not nation-states themselves and thus exploit extraordinary flexibility in adapting to evolving conflicts.

• The capabilities of non-state actors are currently rather limited. They cannot take down the electric grid in the United States, for example, but that will change. We have five or perhaps even 10 years at the outside before the capabilities of non-state actors become almost equivalent to those of nation-state actors. Thus, the window of opportunity to get our strategy right is limited, and the U.S. must take advantage of the time while it can.

• The hardest part of the game is attribution. Knowing who the other side is and what their motivations are is the most difficult challenge of all. How does the U.S. deal with that? Who are these people? What are their true motivations? That is not something that can be fixed technologically. In the end, the U.S. must get better at it, but it is not something for which the same confidence in identifying the enemy can be obtained that is often found in the kinetic world.

The military often talks about “whole of government” approaches to winning wars when “winning” is more than just the battlefield victory over an enemy’s military force. When it comes to cyber warfare, “whole of government” is the only approach that will work against the array of potential adversaries that are exploiting the cyber domain to accomplish their objectives. Integrating military and civilian activities, collecting intelligence, and building a host nation’s security capabilities are all critical elements when combating both state and non-state entities. The full suite of military, intelligence, diplomatic, law enforcement, information, financial, and economic tools will come into play in the new age of cyber warfare.

Organizing for Cyber Warfare

A strategy implies proper organizations and capabilities for fighting a war, but the current manifestations of both are in need of substantial review and investment. During the past several years, many cyber analysts—this author among them—thought the best approach for the U.S. government in dealing with growing cyber threats was to maximize federal government control. What was needed, so the argument went, was a strong cyber czar who had budgetary and directive authority over as much of the government’s cyber capabilities and responsibilities as possible in order to centralize planning for and response to cyber attacks.

Unfortunately, this was precisely the wrong approach to take in dealing with cyber warfare as it has evolved over time. Cyberspace is the world’s most distributive dynamic domain. More than 3.5 billion people and more than a trillion things are connected to the network across the globe. It changes on a daily, even hourly, basis. The advanced, persistent threats that are intruding on Department of Defense (DOD) .mil computers today did not exist six months ago. They are newly
and purposefully built for that enterprise. A centralized hierarchy seems a poor fit for conflict with a diverse, multifaceted, morphing opponent in a battle space that changes every day.

The “big military” complex does a lot of things well, but one of the things it does not do well is turn quickly. The military’s conceptual turning radius is like that of an aircraft carrier, not a Corvette. The military’s major component in dealing with the cyber threat is U.S. Cyber Command (CYBERCOM), a sub-unified command that reports to U.S. Strategic Command. Though it was established only seven years ago in 2009, proposals are already being made to turn it into an independent command.

Given a lengthy pattern of behavior within the Pentagon, it is reasonable to expect that in spite of best efforts to the contrary, CYBERCOM is likely to feature many of the defining characteristics of very large military organizations: lots of rules; lengthy, hierarchical reporting chains; stifling acquisition rules; and a battalion of staff judge advocates (lawyers) who will oversee cyber activities down to the lowest levels of the organization. In this conflict space, however, a model based on “big military” design is the wrong model to pick. Rather, the cyber force needs to be much more akin to special operations: lean, quick to react, and flexible, with a flat administrative structure and possessing the tactical equivalent of a small operational detachment that has top-tier skills and broad authorities to conduct “special mission operations.”

Consider the cyber aspects of some of the recent conflicts America has faced. President Obama continues to consider physical action in Syria or Iraq to confront ISIS. What will ISIS’s cyber response be? What might Syria’s be? The Syrian Electronic Army has already told us that it is going to counterattack if American troops ever go to Syria, and ISIS has threatened to disrupt the American economy. The complexities of conflict are compounded by tactical interdependences and a lack of actionable intelligence.

• What do we know about their capabilities? On the public record, very little—though, to be fair, this may reflect less a gap in our understanding than the existence of capabilities that have not been publicly disclosed. As far as can be seen from the public sources, we do not have anybody on the inside of many of these non-hierarchical organizations.

• What are their likely targets? We may not know, because we do not have any sense of what their capabilities are or any intelligence on their targeting methodologies or what they think are our soft points.

• Do we have targeted weapons that can find the ISIS or Syrian Electronic Army command-and-control servers and take them out without taking offline the entire Syrian and Iraqi electric grids? I suspect that whatever such weapons we have are limited.

• Do we want to take down the entire Syrian and Iraqi electric grid? No, because that is both what the anti-ISIS militia and the Iraqi government are using for their command and control and what the civilians are using to ameliorate the horrible effects of the warfare they are undergoing.

When it comes to the zeroes and ones of DOD efforts to wage cyber warfare, DOD’s organization for battle in cyberspace is typical: offense, defense, functionally focused teams, specified and rigidly envisioned command authorities. DOD speaks of its awareness that “talent” is critical to acquire but hard to find, yet it operates largely within the conventional military model—recruit, train, assign, rotate, and promote—rather than finding and leveraging raw “organic” talent that is optimally suited for this sort of warfare but is very likely not to be found in a conventional military mold.

CYBERCOM has to work trans-domain and trans-COCOM (combatant command), accounting for the nature of the weapons
being used, the diversity and character of actors involved, and the combination of actor interactions. Yet CYBERCOM does not control most of the resources and lacks the authority to dictate to the broad range of largely non-government, private-sector entities that are of critical importance to cyber warfare.

A Separate Command for a Distinct Domain?

One final note: U.S. cyber organization reflects a relatively controversial decision to characterize cyber as a distinct domain. Often, cyber conflict is thought of as a component of information operations (using the cyber domain and related tools to shape perceptions and understanding) or as a subset of electromagnetic warfare (leveraging the same to cause effects on an opponent’s physical ability to conduct operations). Both characterizations are plausible, the first looking at the target area of a conflict (particularly the people in the battle zone) and the latter looking at the cognate physical domain (the assets the people are using to wage war). For this reason, many think that cyber weapons, as a tool of warfare, should be no different from other tools that are incorporated directly into the operational planning of geographic combatant commanders.

The counterargument is that it is useful to characterize the cyber domain as a separate domain, if only because its characteristics are sufficiently different in degree from those of warfare in the kinetic realm that they tend over time to become differences in kind. Under this construct, CYBERCOM is seen as akin to SOCOM (Special Operations Command), managing and employing a unique, highly valued capability that is not defined by region and can be used both for strategic effect and to support conventional military operations of the geographic COCOMs.

Whatever the merits of the debate, the U.S. government has chosen its course. For better or worse, we have characterized the domain based principally on the type of tool (or weapon, if you will) that is used.

But that characterization as a separate command resonates with even greater adverse consequences than a mere category mistake. It seems on reflection to be emblematic of a fundamental misperception of the nature of cyber conflict. To be sure, senior officials often speak of the newness of cyber warfare and acknowledge that new ways of thinking are required, but seven years on, most of the military response to cyber vulnerability reflects, to this author, an inability to reconceptualize military organization and response in light of the domain’s unique characteristics. For example:

- The principal tenet of U.S. legal policy in the domain was a successful effort to adopt existing laws of armed conflict for cyberspace.
- Each of the military services has created within the service a cyber-focused military organization modelled on the fleet/air force model that governs the organization of kinetic military platforms.
- Similarly, CYBERCOM has organized itself along traditional lines with 13 teams, known as Cyber National Mission Teams, responsible for responding to an attack on U.S. critical infrastructure, accompanied by Cyber Combat Mission Teams. To address a lack of training, CYBERCOM has instituted a training system to create “common and strict operating standards” for U.S. cyber operators.

Perhaps this is the right course. To be fair, the Mission Team approach does look somewhat like a special operations approach of the sort this author has advocated. Looking back 10 years from now, we may conclude that these more or less traditional military approaches to conflict in the cyber domain were the right ones.

Nevertheless, one may be skeptical. Considering how cyber capabilities are morphing into a hybrid form of conflict, some of this
seems misguided. Traditional military law, training, procurement, and organization are insufficiently nimble to be responsive to the democratization of conflict in cyberspace. We are seeing a sea-change in the capability of non-state actors, ad hoc groups, and even individuals that allows them to compete on an almost level playing field with nation-states and do significant damage to our national security interests. If we do not reconceptualize how we are thinking about cyber security, cyber policy, and cyber conflict, we are going to miss the boat.

Conclusion

We are facing a new world that is replete with new challenges and rapidly evolving requirements for new ways to respond to those challenges. Anonymous and its ilk are a harbinger: Power and force are being democratized, and we are not ready for it. We are in the midst of a Kuhnian paradigm shift from a time when nation-states had a monopoly on the use of significant force to a time when destructive potential in cyberspace is increasingly available to anyone with the technical skills to employ it anywhere in the world from anywhere in the world irrespective of borders, authorities, or affiliations.

If this is the case, then our current military strategy for operations in cyberspace is focused on the wrong enemy at the wrong time, using the wrong tools and with the wrong hierarchy. This almost certainly means that we are setting ourselves up for catastrophic failure that will lead to nearly unimaginable consequences. Crafting a relevant and effective set of capabilities and response options is therefore a matter of increasing urgency.

The U.S. must get its cyber act together soon: Time is running out.
Endnotes:


Operational Concepts and Military Strength

Antulio J. Echevarria II

What are operational concepts, and how do they contribute to military strength? Essentially, operational concepts are generic schemes of maneuver. They provide the conceptual basis for operational planning and influence the design and employment of military forces. We can think of a military force as a specific slice of military strength. A party’s military strength is, in other words, the aggregate of its military forces. Operational concepts provide a way to convert military strength into military power: the ability to employ military force where and when we want to employ it.

Military power is, of course, relative; it depends as much on our own capabilities as it does on those of our rivals. An Air Force that cannot penetrate an opponent’s air defenses, for example, does not offer much in the way of genuine military power. Operational concepts can tilt the balance (or imbalance) in our favor by creating a functional or employment advantage, and the magnitude of that advantage can mean the difference between success and failure. Operational concepts can be tacit or explicit, planned or emergent. As generic schemes of maneuver, they link “ends” to “means” in military strategy and generally serve as the glue that holds it together.

At the same time, operational concepts have significant downsides. Specifically:

• They usually are poorly defined in military doctrine or shrouded in jargon, which in turn leads to confusion.

• The process by which they are developed is decidedly subjective. Despite many and varied efforts to make that process more objective, it invariably reflects service biases and preferences. That influence can be a virtue or a vice; often, it is a combination of both.

• While operational concepts clearly enable the exercise of military power, they also surely hinder it. This is true mainly because turning an operational concept into doctrine requires a broad and sustained commitment or buy-in, which in turn means opportunity costs in the form of exploring other ideas. This is especially the case with successful concepts such as AirLand Battle, which can breed complacency.

Operational Concepts in Joint Doctrine

The U.S. military’s definition of an operational concept can be found in the Joint Chiefs of Staff’s Joint Publication 1 (JP-1), the current version of which states:

Joint concepts examine military problems and propose solutions describing how the joint force, using military art and science, may
operate to achieve strategic goals within the anticipated future security environment. Joint concepts lead to military capabilities, both non-materiel and materiel, that significantly improve the ability of the joint force to overcome future challenges.¹

Unfortunately, this definition tells us what an operational concept does, not what it is. The failure to define something occurs frequently in U.S. military doctrine and stems from the dogmatic overuse of the active voice and a misplaced aversion to the verb “to be.” It amounts to a failure to communicate that undermines the chief purpose of doctrine, which is to establish a baseline for how the U.S. military operates. Such an understanding benefits not only all of the services, but also our allies and strategic partners. Achieving that purpose requires defining what things are, not just what they do.

Despite these definitional shortcomings, JP-1 does provide useful information about how the U.S. military develops its operational concepts. The purpose of such concepts is to propose “solutions to compelling, real-world challenges both current and envisioned.”² Operational concepts must offer “clear alternative[s]” to existing doctrine or capabilities and “demonstrate evidence of significant operational value relative to the challenges under consideration.” They are to be “idea-focused” and thus not “constrained by existing policies, treaties, laws, or technology.”³ Each concept is to be developed “collaboratively” with the participation of all U.S. military services and evaluated “rigorously” in war games, workshops, and other forums to identify its strengths and weaknesses and to ensure that it actually solves the specified problem.⁴

The evaluation process (Joint Concept Development Process) consists of five phases or steps: prospectus development, concept research and writing, concept evaluation, coordination and approval, and implementation.⁵ Once an operational concept is approved, which can take between 18 and 24 months, it is then fed into the “Joint Force Development Life Cycle.”⁶ The purpose of this cycle is to identify any changes in military doctrine, professional education and training, and equipment required by the new concept. Once operational concepts have passed through the joint development life cycle, they become the overarching “ways” that link “ends” and “means” within the framework of contemporary military strategy.

Today, military strategy is typically thought of in terms of four critical variables: ends or objectives (what we want to achieve); ways or courses of action (how we propose to achieve it); means or resources (what we can reasonably make available); and risk (our assessment of the probability of success).⁷ As generic ways to influence force structure and design, operational concepts can also affect the level of risk, both favorably and unfavorably.

However, there are notable pitfalls in this process. For instance, stripping away political constraints may allow for maximum intellectual creativity, but it also creates an artificial environment wherein policies can be set aside, which in turn leads to operational approaches divorced from the most important kind of real-world challenges: policy constraints. This particular pitfall seems all the more egregious given how the U.S. military’s experiences in Afghanistan and Iraq have revealed the necessity for greater interagency coordination, or a “Whole of Government Approach.”⁸ Would it not be better to acknowledge political realities, perhaps as both constraints and opportunities, at the outset and then develop an operational concept within them and with full interagency participation?

Moreover, while operational concepts can pinpoint the need for new military hardware, they can also be reverse-engineered to justify developing or retaining preferred pieces of equipment or force structure. Because operational concepts influence force structure and military strategy, the stakes are high for each service, which in turn makes cross-service collaboration and objective evaluation that much more difficult. As a consequence, the process of concept development can devolve into a form of horse-trading, with one service...
supporting another in return for an endorsement of its own concept later. The result might be a concept that simply avoids making the hard choices.

An example of a concept that avoided hard choices was Joint Vision 2010 and its successor Joint Vision 2020. It essentially permitted each of the services to continue to develop its own suite of capabilities under the umbrella concept of Full Spectrum Operations. These capabilities—Dominant Maneuver, Precision Engagement, Focused Logistics, Full-Dimensional Protection—put a “mark on the wall” but ultimately meant business as usual for each of the services.

**Operational Concepts in Practice**

Given the vulnerabilities in the Joint Concept Development Process, it should not be surprising that our track record has been mixed. Some concepts, like AirLand Battle, have proved successful; others, such as Effects-Based Operations, have failed; and still others, such as Air-Sea Battle, are under development.

**AirLand Battle.** AirLand Battle, one of the most prominent examples of a successful operational concept, was true to most of the criteria specified in joint doctrine. In 1982, AirLand Battle became the foundation for U.S. military doctrine. It also served as one of the principal “ways” in the West’s military strategy of deterrence during the Cold War, which in turn supported its grand strategy of containment. Although it was never tested against the Warsaw Pact, it was the basis for the operational plan that defeated the Iraqi army in Operation Desert Storm in 1991. AirLand Battle provided a blueprint, a generic scheme of maneuver, for how air and ground forces should operate to stop and ultimately destroy a Soviet-style attack in Central Europe.

The compelling, real-world problem that the concept addressed was how to defeat a numerically superior foe while avoiding a costly war of attrition in a highly lethal environment, particularly one that might include nuclear and chemical weapons. The answer was to put a premium on quality: highly trained troops with better morale, armed with superior weapons, and able to shoot, move, and communicate more efficiently than their foes. Maintaining mobility and a high tempo of operations was essential, as was striking at vital elements beyond the first echelon of the enemy force. Armored and mechanized formations were to block and channel the first echelon of an enemy’s advance, while attack helicopters and fixed wing aircraft were to strike along the enemy’s flanks and concentrate on destroying the command-and-control elements in its second and third echelons.

The key methodological innovation, therefore, was attacking in a synchronized manner throughout the depth of the “extended battle-field.” That, in turn, meant tying the distance between each echelon to the time available to act, all of which was based on a doctrinal template of how the Soviets should attack. Had the Warsaw Pact been able to deviate from that template in any significant way, which was considered highly unlikely, AirLand Battle would have become unhinged, though it might not necessarily have failed outright.

AirLand Battle profoundly influenced the Army’s operational doctrine. It propelled the operational level of war from a matter of debate to an item of doctrine, and it converted Clausewitz’s theory of “center of gravity” and the concentration of superior combat power against it. It also reinforced the need for new land-power requirements: the M1 Abrams Tank, Bradley Fighting Vehicle, Patriot Antiaircraft System, Apache Attack Helicopter, and Blackhawk Utility Helicopter, which became known as the “Big Five.” These systems, it bears noting, were still outnumbered by the Soviets’ “Big 7” (T-72 Tank, BMP Amphibious Assault Vehicle, ZSU-23/4 Anti-Aircraft System, Hind-D Helicopter, 152mm SP Gun, 122mm SP Gun, and SA-3 Surface-to-Air Weapon) but were considered more than a match qualitatively.

AirLand Battle also had the advantage of replacing an unpopular, short-lived, and
perhaps dubious concept called Active Defense.\textsuperscript{22} This concept embraced rather than eschewed attrition—withdrawal just ahead of the Soviet advance, forcing it to deploy, attriting it while it did so, and withdrawing again before becoming decisively engaged. It was less about trading space for time than it was about achieving favorable exchange ratios (better than 3:1) on a relentless basis. As its critics noted, however, it aimed more at avoiding defeat than winning in a manner that might give political leaders something to bargain with at the negotiating table.

Collaboration between the U.S. Army and U.S. Air Force in the development of AirLand Battle was extensive, if fraught with friction.\textsuperscript{23} The U.S. Navy was involved only tangentially, since it already had a major mission, detailed in the 1986 \textit{Maritime Strategy}: to protect sea lines of communication and supply across the Atlantic Ocean and the Mediterranean Sea, to provide supporting air cover where possible over Western and Central Europe, and to maintain the ability of its submarine fleets and carrier battle groups to strike targets inside the Soviet Union.\textsuperscript{24} The Navy’s mission clearly supported deterrence in Western Europe and containment, and because its service equities were not threatened, it had no reason to obstruct the development of AirLand Battle.\textsuperscript{25}

AirLand Battle was not without its opportunity costs. Those came in the form of “military operations other than war” (MOOTW, or missions ranging from shows of force to humanitarian assistance), which were treated as “lesser includeds.”\textsuperscript{26} However, not all such operations could be treated as miniature AirLand Battles. Some examples were the interventions in El Salvador (1979–1991) and Colombia (1978–2011); the aborted rescue operation in Iran (1980); the interventions in Grenada (1983) and Panama (1989); and the humanitarian assistance operation in Somalia (1992–1994). From this sample, the United States might claim four “wins” and two “losses,” or a 66 percent success rate—simply not good enough.\textsuperscript{27}

**Effects-Based Operations.** In contrast to AirLand Battle, Effects-Based Operations (EBO) did threaten service equities: specifically, those of the Army and Marine Corps. EBO was officially defined as a “process for obtaining a desired strategic outcome or ‘effect’ on the enemy through the synergistic, multiplicative, and cumulative application of the full range of military and other national capabilities at the tactical, operational, and strategic levels.”\textsuperscript{28} In short, it was to afford policymakers a menu of “effects” from which they might choose the one they desired.

EBO belonged to an umbrella concept referred to as Network-Centric Warfare, credit for which belongs chiefly to the late Admiral Arthur K. Cebrowski of the U.S. Office of Force Transformation.\textsuperscript{29} It did not respond to a specific real-world challenge, but rather attempted to leverage information technology in a manner that would make warfare more precise, less costly, and ultimately more useful as an instrument of policy.

Coalition forces attempted a version of EBO during the Kosovo conflict in 1999 and in the early stages of the campaigns in Afghanistan and Iraq. While destruction of matériel and disruption of infrastructure and communications were readily accomplished, effects beyond these accomplishments remained elusive. In 2008, the U.S. Joint Forces Commander, U.S. Marine Corps General James Mattis, shelved the concept for being inimical to war’s unpredictable nature.\textsuperscript{30} By then, however, EBO and NATO’s counterpart EBAO (Effects-Based Approach to Operations) were already integrated into several nations’ operational doctrines.

As happened with AirLand Battle, the West’s experiments with EBO led to significant opportunity costs in terms of exploring other concepts. In theory, EBO could be employed broadly across the diplomatic, informational, military, and economic (DIME) dimensions of national power; in practice, it was applied only to a narrow segment of the spectrum of operations, a segment in which the U.S. military already excelled. The other
agencies within the U.S. government failed to embrace it.

Consequently, EBO amounted to a refinement of military operations in a single portion of the spectrum of conflict; not unlike AirLand Battle, it proved ill-suited to humanitarian assistance or similar operations requiring physical control and human presence and interaction: in other words, shoes as well as boots on the ground.\textsuperscript{31} Put differently, if the post–Cold War security environment was really characterized by unprecedented uncertainty, as many claimed, it would have been wiser to develop a broad array of capabilities and ways of thinking to avoid what historian Sir Michael Howard famously referred to as “being too badly wrong.”\textsuperscript{32}

**Air-Sea Battle.** Although EBO was shelved, it was by no means dead. Its principles resurfaced in the concept of Air-Sea Battle, which was unveiled (perhaps prematurely) in 2010. Air-Sea Battle generated controversy almost immediately, but it did respond to a specific real-world challenge. The version unveiled in 2010 was a “point-of-departure” concept designed to address China’s growing anti-access/area denial (A2/AD) capabilities along the Pacific Rim. As its authors explained:

> These capabilities threaten to make US power projection increasingly risky and, in some cases and contexts, prohibitively costly. If this occurs, the United States will find itself effectively locked out of an area that has been declared a vital strategic interest by every administration for the last sixty years.\textsuperscript{33}

The U.S. military already had a doctrine for conducting “forcible entry” operations, which was barely two years old, but it applied mainly to the kinetic use of force in time of war.\textsuperscript{34} It did not address the larger strategic goal of maintaining a “favorable conventional military balance throughout the Western Pacific region” with the ability to “deter China from acts of aggression or coercion in the region.”\textsuperscript{35} Thus, the problem posed by the People’s Liberation Army’s growing A2/AD capabilities was (and remains) a compelling real-world challenge worthy of a revised operational concept—provided that concept also addresses how to augment military capabilities with other forms of power in order to gain more deterrence value. The unclassified versions of Air-Sea Battle have not yet addressed this issue.

In addition, several failures related to insular thinking and timing undercut Air-Sea Battle. The concept’s authors did not adequately incorporate Army and Marine Corps equities into its development. That \textit{faux pas} was later corrected, at least partially, when Air-Sea Battle was subordinated to the Joint Operational Access Concept (JOAC), which took a more service-integrated approach to solving the access problem.\textsuperscript{36} Nonetheless, it was an egregious error of omission at a time when rumors of significant downsizing across the Department of Defense (DOD) were gaining momentum.

The lack of full cross-service integration led senior Army and Marine Corps leaders to believe that their services were to be the “bill-payers” for the “Pacific Rebalance” and for implementing Air-Sea Battle. It is little wonder, then, that the concept was greeted with such hostility.

Second, although its authors took pains to explain that Air-Sea Battle was not about “containing” or “rolling back” China, but rather about “offsetting the PLA’s unprovoked and unwarranted military buildup,” it did not play that way in the media.\textsuperscript{37} Critics reacted sharply, claiming that Air-Sea Battle was a poor substitute for a military strategy (which, however, it was not intended to be) and that it would likely provoke China precisely when the United States wanted to avoid doing so. As official documents tried to make clear, Air-Sea Battle was not intended to function in isolation, but to be combined with “security assistance programs, and other whole-of-government efforts.”\textsuperscript{38} It signaled a commitment by the United States to maintain an “escalation advantage” in conflict while sustaining “security and prosperity” in peacetime.\textsuperscript{39}

The central idea of Air-Sea Battle in its unclassified form is “to develop networked,
integrated forces capable of attack-in-depth to disrupt, destroy, and defeat adversary forces.” In this regard, it shows the influence of network-centric operations, a concept first advanced in the 1990s as part of a DOD-wide effort to capitalize on the revolution in military affairs (RMA).

- A “networked” force is one in which command and control can be exercised instantaneously across service-specific barriers or protocols not only through technological means, but also through “habitual relationships across service, component, and domain lines.”

- The notion of an “integrated” force goes beyond the traditional idea of task-organizing for a mission; instead, units are to be “pre-integrated” with regard to joint and combined training and procedures well before arriving in theater and, ideally, in terms of material management, thereby ensuring interoperability and minimal redundancy.

- The ability to “attack-in-depth” refers to the use of kinetic and non-kinetic means in the form of offensive and defensive fire and movement to accomplish one of three outcomes or some combination of them: disrupting an adversary’s “effects chains” (the opponent’s process of finding, fixing, tracking, targeting, engaging, and assessing) by impeding command and control and the flow of information; destroying A2/AD platforms and systems; and defeating weapons and formations “post-launch.” Attack-in-depth thus reflects the influence of the ideas that underpinned EBO, though the term itself is avoided.

In fairness, Air-Sea Battle was exactly what it claimed to be: a single answer to a specific operational challenge. While that challenge is not new, the relentless advance of technology is making it more difficult. The concept placed a very high, perhaps idealistic “mark on the wall” with regard to the level of capabilities and competencies necessary to execute it. It is still under development as part of the Joint Operational Access Implementation Plan.

In the interim, the JOAC serves as the doctrinal concept for the U.S. military’s working solution to the contemporary A2/AD challenge. In brief, the JOAC says we can project force in an A2/AD environment by using “cross-domain synergy” to achieve superiority in specific domains, which will then lead to a certain amount of “freedom of action.” Interestingly, the tone is reminiscent of the optimism that characterized military theory on the eve of World War I, which proposed using firepower superiority to overcome the anticipated strength of the defense.

**Emergent Concepts**

Some operational concepts are emergent. These concepts develop not in anticipation of future problems, but as responses to challenges that arise during a conflict.

An example occurred most recently in the campaigns in Iraq and Afghanistan with the emergence of U.S. counterinsurgency doctrine. The doctrine was not new; rather, it was a rediscovery of previously accepted principles. Both the Army and Marine Corps already had a substantial number of official publications addressing guerrilla warfare and insurgencies, but that doctrine had all but faded from institutional memory, partly because of the residual influence of AirLand Battle and partly because of the enthusiasm with which the Office of the Secretary of Defense pushed its technology-based transformation program in the 1990s. It thus had to be rediscovered and updated.

When enemy fighters shifted to insurgent techniques, therefore, many Coalition formations had to adapt without the benefit of either explicit or tacit operational concepts. Nonetheless, some American units were employing counterinsurgency techniques by 2004 and 2005, well before official U.S. counterinsurgency doctrine appeared. Several
scholars described this adaptation as a revolution from the top down, while others portrayed it as coming from the “bottom-up.”

In truth, it was neither. The emergence of counterinsurgency techniques came into play more or less from a “sideways” direction, or laterally, through mid-level officers and noncommissioned officers who exercised reach-back capabilities and consulted with civilian experts and with each other to exchange information and share knowledge about what worked and what did not. Many counterinsurgency principles and practices (as well as healthy criticism of them) emerged through sheer trial and error and through the common sense (or experienced judgment) of brigade and battalion commanders.

Official U.S. counterinsurgency doctrine, when it did appear, helped to codify and standardize—that is, render explicit—many of the procedures that were already in play, albeit unevenly and perhaps even poorly in some cases, and augmented them with others. It also situated such practices within a generic scheme of maneuver, which in turn rationalized them. The various stances in the counterinsurgency debate are well known and need not be addressed here.

The U.S. military’s rediscovery of counterinsurgency techniques was part of the process of adaptation that occurs relentlessly in wartime. Adaptation is simply how we cope with a situation or an adversary; in contrast, innovation is how we overcome one or the other—or, in some instances, both.

Conclusion

As we have seen, operational concepts are integral to military strength. They help to convert potential military strength into military power, an unquestionably essential function. However, they also have significant downsides. In part, these downsides stem from the processes by which operational concepts are developed. As JP-1 revealed, operational concepts are to be developed in a manner that affords a maximum amount of intellectual creativity. Paradoxically, this approach is also what makes operational concepts—whether AirLand Battle, Effects-Based Operations, Air-Sea Battle, or counterinsurgency doctrine—vulnerable.

In theory, each service should know best what it needs to be able to operate in the future security environment. In practice, however, what the services know is sometimes exquisitely irrelevant to the needs of policymakers. Armed conflict can have the effect of forcing policymakers and military professionals outside of their comfort zones. That, in short, is what led to the emergence of counterinsurgency as an operational concept; it was an answer of sorts, however flawed, to a situation that the concept development process, and all of its attendant evaluation and war-gaming, ought to have anticipated and yet did not.

The evaluation part of the process ought to force political and military leaders outside of their comfort zones long before the fighting starts. Otherwise, we are engaging in a tautology in which our operational concepts are designed to fight the abstract battles we like instead of the real wars we do not like. The bitter irony is that sometimes the tautology works. Operation Iraqi Freedom was the real war that suited the abstract battle. We would do well to remember, though, that such victories will offer little comfort when the opportunity costs of our tautology come due.

What about the future? Events in Eastern Europe and East Asia suggest that there is a need for an operational concept capable of exerting better deterrent and coercive leverage. Might we see some form of an intellectual blend—a maneuver-oriented concept that can coerce, married to an A2/AD concept that can deter? Certainly, the real-world challenge is there.
Endnotes:


2. Ibid., p. VI-9.

3. Ibid.

4. Ibid., p. VI-10.


11. Ibid.


16. Although the European theater was the concept’s primary focus, AirLand Battle could also apply to the defense of South Korea, albeit with some modifications due to differences in terrain and enemy capabilities.


27. Further qualifications are possible: The U.S. intervention in El Salvador (55 advisers, $6 billion) did not bring victory for the Salvadoran armed forces, but it did prevent their defeat; America’s assistance to Colombia since 2000 (800 soldiers, 600 contractors, $7 billion) coincided with the success of President Alvaro Uribe’s strategy. See Antulio J. Echevarria II, *Reconsidering the American Way of War: US Military Practice from the Revolution to Afghanistan* (Washington: Georgetown University Press, 2014), and Stephen Watts, Caroline Baxter, Molly Dunigan, and Christopher Rizzi, *The Uses and Limits of Small-Scale Military Interventions* (Santa Monica, CA: RAND Corporation, 2012).


35. Van Tol et al., *Air Sea Battle*, p. xi.


37. Van Tol et al., *Air Sea Battle*, p. x.


39. Ibid.

40. Ibid., p. 4.

41. Ibid., p. 6.

42. Ibid.


On Strategy and Strategic Planning: Repairing America’s Strategic “Black Hole”

Mackubin Thomas Owens

Strategy has long been the subject of scholarly study and policy analysis. Historians and social scientists alike have written widely about strategic thought, process, and practice. Scholars continue to dissect the meaning of strategy. War colleges teach courses on the subject, as do civilian colleges. Yale University, for instance, has a well-regarded program on grand strategy, and other universities have followed suit.

Strategy and strategy-making are complex phenomena, not reducible to a simplistic mechanical process, and the making of strategy deserves more study than it often receives. In many respects, U.S. strategic planning has been rendered nearly useless because the processes have become routinized and thereby trivialized. Legislatively required documents such as the National Security Strategy and the Quadrennial Defense Review (QDR) initially may have been useful but now are merely periodic bureaucratic exercises.

The result is what Colin Gray calls “a black hole where American strategy ought to reside.” What the United States needs is a return to the long-range strategic planning process that it implemented during the Cold War.

On Strategy and Policy

When all is said and done, strategy is ultimately best understood as the interaction of three things, all within the context of risk assessment:

- **Ends** (the goals or objectives that the strategic actor seeks to achieve);
- **Means** (the resources available to the strategic actor); and
- **Ways** (the strategic actor’s plan of action for utilizing the means available).

In essence, any strategy worth the name should articulate a clear set of achievable goals; identify concrete threats to those goals; and then, given available resources, recommend the employment of specific instruments to meet and overcome those threats.

A good strategy also seeks to minimize risk by, to the extent possible, avoiding mismatches between strategy and related factors. For instance, strategy must be appropriate to the ends as established by policy. Strategy also requires the appropriate tactical instrument to implement it. Finally, the forces required to implement a strategy must be funded, or else it must be revised. If the risk generated by such policy/strategy, strategy/force, and force/budget mismatches cannot be managed, the variables must be brought into better alignment.

History clearly teaches that the development of a coherent strategy is absolutely essential to national security in times of both war and peace. In the absence of a coherent strategy, non-strategic factors such as bureaucratic and organizational imperatives and the
vicissitudes of domestic politics will fill the void to the detriment of national security.

Modern strategic studies can be said to begin with the division of the art of war into the theory of “the use of engagements for the object of the war” (strategy) and “the use of armed forces in the engagement” (tactics) by the great interpreters of Napoleonic warfare, Baron Antoine Henri de Jomini and Carl von Clausewitz. As the latter wrote:

Strategy is the use of the engagement for the purpose of the war. The strategist must therefore define an aim for the entire operational side of the war that will be in accordance with its purpose. In other words, he will draft the plan of the war, and the aim will determine the series of actions intended to achieve it: in fact, shape the individual campaign and, within these, decide on the individual engagements.

These 19th century writers originated the modern conception of strategy as the art of assembling and employing military forces in time and space to achieve the goals of a war. While such writers normally limited their use of “strategy” to mean the application of military forces to fulfill the ends of policy, it is increasingly the practice today to employ the term more broadly so that one can speak of levels of strategy during both peace and war. Accordingly, more often than not, strategy now refers not only to the direct application of military force in wartime, but also to the use of all aspects of national power during peacetime to deter war and, if deterrence fails, win the resulting conflict.

This more expansive usage of strategy inevitably overlaps with the common meaning of “policy,” which is defined as the general overall goals and acceptable procedures that a nation might follow and the course of action selected from among alternatives in light of given conditions. In their military history of the United States, Allan Millett and Peter Maslowski define defense policy as “the sum of the assumptions, plans, programs, and actions taken by the citizens of the United States, principally through governmental action, to ensure the physical security of their lives, property, and way of life from external military attack and domestic insurrection.” For our purposes, “policy” refers primarily to such broad national goals as interests and objectives, and “strategy” to the alternative courses of actions designed to achieve those goals, within the constraints set by material factors and geography.

In general, strategy provides a conceptual link between national ends and scarce resources, both the transformation of those resources into means during peacetime and the application of those means during war. As such, it serves three purposes.

- Strategy relates ends or the goals of policy (interests and objectives) to the limited means available to achieve them. Both strategy and economics are concerned with the application of scarce means to achieve certain goals, but strategy implies an adversary who actively opposes the achievement of the ends.

- Strategy contributes to clarification of the ends of policy by helping to establish priorities in the light of constrained resources. In the absence of established priorities among competing ends, all interests and all threats will appear equal. In the absence of strategy, planners will find themselves in the situation described by Frederick the Great: “He who attempts to defend too much defends nothing.”

- Strategy conceptualizes resources as a means in support of policy. Resources are not means until strategy provides some understanding of how they will be organized and employed. Defense budgets and manpower are resources. Strategy organizes these resources into divisions, wings, and fleets and then employs them to deter war or to prevail should deterrence fail.

The first two functions make it clear that a broad national strategy must shape strategies for various regions and theaters by
prioritizing them. In terms of warfighting, the national strategy establishes the desired goals in a theater, linking operational considerations to the requirements established by national authorities. Based on guidance from higher authorities, the theater commander determines the desired outcome within his area of responsibility. The staff then develops war plans based on an array of plausible scenarios. Using various force planning models and war games to determine force size and mix, the theater commander’s staff then derives the force necessary at the outset of a campaign to achieve the desired outcome.

In addition to determining the required force, staffs at all levels also determine the schedule for deploying forces from out of theater. Part of this determination is establishment of the Time-Phased Force Deployment Line, designating in a detailed manner the timeline for forces to be deployed to the theater. The higher-level strategies also establish priorities among the various theaters, indicating which will be the site of the main effort and which might be designated “economy of force” in the event that crises occur in more than one theater simultaneously.

National strategy thus guides “force apportionment,” the distribution of existing forces among the various theaters. During World War II, national strategy dictated a policy of “Europe first.” During the Cold War, U.S. strategy dictated a focus on Europe followed by the Asia–Pacific and finally by the Greater Middle East.

Of course, warfighting and war planning are only part of the theater commander’s job. He is also responsible for shaping the theater in hopes of advancing U.S. interests without recourse to war, engaging the governments within the region and developing the necessary security infrastructure to maintain a favorable state of affairs. In this regard, the theater commander employs such tools as security assistance, military exercises, and humanitarian support. The theater commander’s actions are not strictly military in nature; diplomacy and interagency operations play a major role in the development and implementation of each geographic command’s Theater Security Co-operation Plan.

The final function of strategy is to serve as a guide to force planning. In theory, the strategy–force planning process is logical. The planner first identifies national interests and the objectives required to achieve those interests. The planner then conducts a net assessment in order to determine the ability of adversaries to threaten those interests or to interfere with the achievement of national objectives. These represent the “operational challenges” that U.S. forces must surmount in order to implement the strategy. Next, the planner forges a strategy to overcome operational challenges and a budget to fund the capabilities and operational concepts that are needed to implement the strategy.

The execution of any chosen strategy requires the fulfillment of certain strategic requirements. These requirements determine the necessary military capabilities and operational concepts, which in turn drive the acquisition of forces and equipment. Thus, if there is a strategic requirement for a particular capability, the forces or equipment needed to provide that capability presumably should be obtained. To overcome these operational challenges and confront plausible future areas of military competition, the United States must develop new operational concepts.9

Although strategy can be described as the conceptual link between ends and means, it cannot be reduced to a mere mechanical exercise. Instead, it is “a process, a constant adaptation to shifting conditions and circumstances in a world where chance, uncertainty, and ambiguity dominate.”10 It is a mistake to attempt to reduce strategy to a single aspect, although it is not unusual for writers on strategy to try.11 Clausewitz dismissed as simplistic the reduction of strategy to “principles, rules, or even systems” because, on the contrary, strategy “involves human passions, values, and beliefs, few of which are quantifiable.”12

Strategy, properly understood, is a complex phenomenon comprising a number of
elements. Among the most important of these are geography; history; the nature of the political regime, including such elements as religion, ideology, culture, and political and military institutions; and economic and technological factors. Accordingly, strategy can be said to constitute a continual dialogue between policy on the one hand and these various factors on the other.

### Strategy as a Dialogue Between Policy and National Power

To be successful, strategy-making must be an interactive process that takes account of the interplay of all relevant factors. An inflexible strategy may be worse than no strategy at all, as the Germans discovered in 1914 and the French found in 1940. To paraphrase Gray, strategy is the product of the dialogue between policy and national power in the context of the overall international security environment.

Real strategy must take account of such factors as technology, the availability of resources, and geopolitical realities. The strategy of a state is not self-correcting. If conditions change, policymakers must be able to discern these changes and modify the nation’s strategy and strategic goals accordingly.

When strategy-makers do not adapt to changing conditions, serious problems can result. Jakub Grygiel shows how a failure to adapt strategy to geopolitical change led to the decline of Venice (1000–1600); the Ottoman Empire (1300–1699); and Ming China (1364–1644). Each actor faced changing circumstances but made wrong strategic choices. These cases are cautionary for the United States, since it has faced substantial geopolitical changes of great magnitude since the end of the Cold War: the decline and then reassertion of Russian power, the expansion of terrorist organizations, the rise of China, disorder in the Greater Middle East, and the new geopolitics of energy.

### Strategic Culture

Another important aspect of strategy-making is the “strategic culture” of a state or nation. By applying the notion of strategic culture, analysts attempt to explain continuity and change in national security policies, thereby creating a framework that can explain why certain policy options are pursued by states that share a given strategic culture.

For instance, historians have noted that the strategic culture of sea powers tends to differ from the strategic culture of land powers. Thus, one sees similarities between the strategic approaches of Athens, Great Britain, and the United States on the one hand as opposed to the strategic approaches of Sparta, Germany, and Russia on the other. China seems to possess a discernible strategic culture traceable to Sun Tzu and other Chinese military thinkers.

According to Kerry Longhurst:

>[A] strategic culture is a distinctive body of beliefs, attitudes and practices regarding the use of force, which are held by a collective and arise gradually over time, through a unique protracted historical process. A strategic culture is persistent over time, tending to outlast the era of its original inception, although it is not a permanent or static feature. It is shaped and influenced by formative periods and can alter, either fundamentally or piecemeal, at critical junctures in that collective’s experiences.

For Carnes Lord, strategic culture constitutes the traditional practices and habits of thought by which military force is organized and employed by a society in the service of its political goals.

One of the charges often brought against American strategic culture is that it confuses technological superiority with strategy itself. For instance, critics of the efforts to
“transform” the U.S. military in the early years of the 21st century claimed that America tends to seek technological fixes for strategic problems in an attempt to remove itself from the sharp end of war.24

Strategy vs. Nonstrategic Factors

In any case, strategy is an indispensable element of national security. Without it, something else will fill the void. For example, in wartime, service doctrines will dominate the conduct of operations if strategy is absent. This state of affairs is captured by Andrew Krepinevich in his characterization of the Vietnam War as “a strategy of tactics.”25

In peacetime, defense planning is usually dominated by domestic policy considerations such as organizational imperatives and congressional politics. In his 1961 book The Common Defense, Samuel Huntington observed that military policy exists in two worlds: the world of international politics and the world of domestic politics. The first world is shaped by such factors as balance of power, wars and alliances, and the use of force and diplomacy to influence the actions of other states in the international arena. The principal “currency” of this arena is “power,” primarily military power. The second world is shaped by interest groups, corporate interest groups, political parties, social classes, and the like. The currency here is the resources provided by society, personnel, money, and matériel.

Military decisions influence and are influenced by both worlds, and a decision in one currency is payable in the other. Huntington called the decisions in the currency of international politics strategic in character. Decisions in the currency of domestic politics are structural. Unless there is a strong and coherent strategic vision to guide defense decisions even during peacetime, defense decision-making is likely to be dominated by structural decisions.26

Levels of Strategy

War and conflict can be divided into several levels. As noted, Clausewitz distinguished between tactics, “the use of armed forces in the engagement,” and strategy, “the use of engagements for the object of war.” It is now common to speak of an intermediate level between strategy and tactics: the “operational level of war,” a realm concerned with the planning and conduct of campaigns to achieve strategic goals within a theater of war.27 The central focus of this essay is the strategic level of war and conflict, which in itself is subject to further subdivision.28

In its broadest sense, strategy is grand strategy. In the words of Edward Mead Earle:

[5]trategy is the art of controlling and utilizing the resources of a nation—or a coalition of nations—including its armed forces, to the end that its vital interests shall be effectively promoted and secured against enemies, actual, potential, or merely presumed. The highest type of strategy—sometimes called grand strategy—is that which so integrates the policies and armaments of the nation that resort to war is either rendered unnecessary or is undertaken with the maximum chance of victory.29

Thus, grand strategy is intimately linked to national policy in that it is designed to bring to bear all the elements of national power—military, economic, and diplomatic—in order to secure the nation’s interests and objectives. Grand strategy can also refer to a nation’s overarching approach to international affairs: isolationism or disengagement, cooperative or collective security, selective engagement, and primacy.30

Finally, grand strategy can allude to a geopolitical orientation: “continental” or “maritime.”31 Whichever meaning is emphasized, the choice of a grand strategy has a major impact on the other levels of strategy and force structure.

Military power is one instrument of grand strategy. How military power is employed in both war and peace is the province of military strategy. In peacetime, military strategy provides a guide to what Samuel Huntington calls “program decisions” and “posturing.” Program decisions involve the strength of military
forces, their composition and readiness, and the number, type, and rate of development of weapons. Posturing is defined by how military forces are deployed during peacetime to deter war (Clausewitz’s “preparation for war”). In wartime, military strategy guides the employment of military force in pursuit of victory (Clausewitz’s “war proper”).

**U.S. Strategic Planning and the Strategic “Black Hole”**

Given the relatively secure position of the United States at least after the War of 1812, the early American national security apparatus—the State Department, War Department, and Navy Department—remained small and primitive compared to those of the European states. Nonetheless, the United States in fact pursued a consistent grand strategy from the Founding until the outbreak of World War II. The objective of this grand strategy—often mistaken for isolationism—was to maintain the security of the United States by means of skillful diplomacy combined with preemption and unilateralism.

With the outbreak of World War II, the requirements of fighting a global conflict in conjunction with allies impelled the United States to develop the sort of national security apparatus we see today, but it was not until the Cold War, the National Security Act of 1947, and subsequent amendments that this structure came of age.

The problem today is that the documents that supposedly inform U.S. strategy do no such thing. They are, at best, pro forma bureaucratic exercises. For instance, the National Security Strategy (NSS), required by the Goldwater–Nichols Act of 1986, presumably serves as the U.S. grand strategy document, defining U.S. security interests, objectives, and goals and providing guidance to those who are charged with executing that strategy. But while there have been some excellent examples in the past, the NSS has lately become little more than a list of aspirations with no real strategic plan for achieving its stated goals.

Other documents intended to supplement the NSS—the National Defense Strategy, National Military Strategy, and Quadrennial Defense Review—have served only to confuse strategic planning. This is especially true of the QDR, which has long been little more than a bureaucratic budgetary exercise that the services “game” in order to protect or expand their shares of the defense budget. In addition, the QDR has recently been required to address the latest fashionable issues of the day, such as “climate change.”

In short, the United States has failed to provide useful strategic guidance for translating national policy into theater strategy and force employment, shaping force structure, and integrating and synchronizing the planning and activities of the Joint Staff, combatant commands, the services, and combat support agencies. As Michele Flournoy and Shawn Brimley have observed:

The U.S. government currently lacks both the incentives and the capacity to support strategic thinking and long-range planning in the national security arena. While individuals on the National Security Council (NSC) staff may develop planning documents for their respective issues, the NSC staff lacks adequate capacity to conduct integrated long-range planning for the president. While some capacity for strategic planning exists in the Department of Defense (DOD), no other department devotes substantial resources to planning for the long-term future. Although the State Department’s policy planning office develops a “big picture” approach in specific policy areas, such as North Atlantic Treaty Organization enlargement or relations with China, it tends (with some exceptions) to focus on issues already on the policy agenda rather than challenges that might loom over the horizon. Nor does it address the types of capabilities the United States should seek to develop to deal with future challenges.

The result is Colin Gray’s strategic “black hole.”

**A Return to Strategic Planning**

Colin Dueck has offered a useful critique of what currently passes for strategic planning.
In particular, he criticizes the centralization of foreign policy planning in the White House under President Obama. He offers six suggestions to correct the problem:

- Develop and execute a meaningful national security strategy early on.
- Restore a proper balance of responsibilities between the NSC and line departments and agencies.
- Appoint a strong national security advisor to play the role of genuine honest broker, policy entrepreneur, and presidential agent.
- Appoint and empower a strategic planning directorate on the NSC staff.
- Create an effective strategic planning board.
- Learn from private[-]sector experience.³⁶

It would also be useful to revisit the U.S. strategic planning approach during the Cold War. Two of the most important documents shaping early Cold War policy and strategy were NSC-20/4, “U.S. Objectives with Respect to the USSR to Counter Soviet Threats to U.S. Security,” signed by President Harry Truman in 1948, and NSC-68, “United States Objectives and Programs for National Security,” signed by President Truman in 1950. Both documents argued for a policy of “containment” against the Soviet Union, the purpose of which was to prevent Soviet expansionism and guarantee the security of America and its allies.³⁷ NSC-68 in particular served as the foundation of the U.S. approach to the Soviet Union until its collapse in the late 1980s.

However, the cost associated with NSC-68 was high: It called for a tripling of the defense budget to build up U.S. conventional forces and to develop a thermonuclear weapon. Concerned about the high cost of defense, President Dwight Eisenhower sought a way to examine existing American containment policy and compare alternative policy options. He settled on a systematic policy exercise to review U.S. foreign policy objectives and recommend a course of action. The exercise, called “Project Solarium,”³⁸ pitted three teams against each other.

- Team A would make the best possible argument for the existing policy of containment, seeking to prevent Soviet expansion in Europe while minimizing the risk of general war.
- Team B would accept containment as a viable policy but be less tentative about its implementation. It would assert that any Soviet or Soviet-sponsored aggression would lead to general war and threaten massive U.S. and allied retaliation using any means necessary.
- Team C would argue for “rollback,” meaning a policy to halt and then reverse Soviet efforts to hold territory by the presence of the Red Army.³⁹

Five weeks later, the teams reconvened and presented their findings to the President.

- Team A argued that the U.S. should develop and implement a more dynamic campaign of political and psychological action against the Soviets. The group rejected any strategy that based its arguments on the acceptance of a risk of general war and recommended “waging peace” with U.S. power by emphasizing the importance of negotiations. It also sought to prevent the use of an active military threat from driving national security strategy even though it gave the concept of force an important role to play—primarily the role of augmenting diplomatic, economic, and political initiatives.
- Team B warned about the rigid nature of “drawing a line,” implying that it could actually increase the risk of war through
inflexibility, but argued that a preponderant show of U.S. force combined with a definitive geographical boundary line could lead to a change in Soviet policy and/or a mellowing of the overall regime. Team B further explained that the allies would not readily accept where to draw the line and that this strategy would be extremely expensive. However, it made the case that the external threat to the U.S. ultimately outweighed the threat to domestic economic stability.

- Team C argued that mere containment was flawed because it had no endgame and let the Soviets read American inaction as fear and acquiescence. It acknowledged that the benefits of “rollback” were speculative but claimed that political and military actions short of general war (for example, covert operations and economic pressure) would be an effective way to take back regions from the Soviet area of control until, ultimately, the Soviet Union changed. Therefore, the U.S. must first put indirect pressure on the Soviet Union by engaging its satellite states and then direct pressure on the Soviet Union itself.\(^40\)

After listening to the presentations, President Eisenhower summarized the arguments of the three teams and opted for the course of action recommended by Team A, which served as the foundation of NSC-162/2, “Basic National Security Policy,” signed by Eisenhower on October 30, 1953. As one commentator notes:

> As Eisenhower observed, “The plans are nothing, but the planning is everything.”\(^42\)

**Conclusion**

Strategy is designed to secure national interests and to attain the objectives of national policy by the application of force or threat of force. Strategy is dynamic, changing as the factors that influence it change. Strategic requirements continue to evolve.

The evolution of strategy over the past 50 years illuminates the interrelationship of ends, means, and the security environment. Potential mismatches between ends and means create risks. If the risks resulting from an ends–means mismatch cannot be managed, ends must be reevaluated and scaled back, means must be increased, or the strategy must be adjusted.

Strategy-making is a central component of defense policy. Without a coherent, rational strategy to guide the development and employment of forces, structural factors such as bureaucratic and organizational imperatives will dominate the allocation of resources for defense, leading to a suboptimal result.

Good strategy requires an effective strategic planning process. Unfortunately, U.S. strategic planning is defective. As a result, U.S. actions against China, Iran, Russia, ISIS, and the like are uncoordinated and incoherent. To advance its national interests in a dangerous and uncertain world, the United States must restore strategic planning and the idea of strategy as a guide to action to a central role.

Strategic planning must look beyond the next budget cycle in order to address the wide array of international challenges the United States faces and advance long-term U.S. interests. The best strategic planning process incorporates both constructive competition and creative cooperation in order to reconcile diverging perspectives. Otherwise, the U.S. strategic black hole will persist.

---

\(^40\) Team C argued that mere containment was flawed because it had no endgame and let the Soviets read American inaction as fear and acquiescence. It acknowledged that the benefits of “rollback” were speculative but claimed that political and military actions short of general war (for example, covert operations and economic pressure) would be an effective way to take back regions from the Soviet area of control until, ultimately, the Soviet Union changed. Therefore, the U.S. must first put indirect pressure on the Soviet Union by engaging its satellite states and then direct pressure on the Soviet Union itself.

\(^41\) As Eisenhower observed, “The plans are nothing, but the planning is everything.”

\(^42\) Strategy-making is a central component of defense policy. Without a coherent, rational strategy to guide the development and employment of forces, structural factors such as bureaucratic and organizational imperatives will dominate the allocation of resources for defense, leading to a suboptimal result.

Good strategy requires an effective strategic planning process. Unfortunately, U.S. strategic planning is defective. As a result, U.S. actions against China, Iran, Russia, ISIS, and the like are uncoordinated and incoherent. To advance its national interests in a dangerous and uncertain world, the United States must restore strategic planning and the idea of strategy as a guide to action to a central role.

Strategic planning must look beyond the next budget cycle in order to address the wide array of international challenges the United States faces and advance long-term U.S. interests. The best strategic planning process incorporates both constructive competition and creative cooperation in order to reconcile diverging perspectives. Otherwise, the U.S. strategic black hole will persist.
Endnotes:
4. Ibid., p. 177.
8. I am indebted to Dr. Robert S. Wood, former dean of the Center for Naval Warfare Studies, Naval War College, for this formulation.
11. For instance, in Strategy, Luttwak reduces strategy to a manifestation of “paradoxical logic.” In his monumental The Art of War in the Western World (Urbana: University of Illinois Press, 1987), Archer Jones reduces strategy to a choice between “persisting” and “raiding.” In World Politics and the Evolution of War (Baltimore: Johns Hopkins University Press, 1995), John Weltman explicates warfare in terms of a simplistic version of Delbrück’s distinction between strategies of annihilation and attrition.
38. Named for the room in the White House where the President issued the directive for the study.
40. Ibid.
41. Ibid.