2001

Summer 2001 Review

The U.S. Naval War College

Follow this and additional works at: https://digital-commons.usnwc.edu/nwc-review

Recommended Citation

This Full Issue is brought to you for free and open access by the Journals at U.S. Naval War College Digital Commons. It has been accepted for inclusion in Naval War College Review by an authorized editor of U.S. Naval War College Digital Commons. For more information, please contact repository.inquiries@usnwc.edu.
In one of his best-known presentations, “The Future Security Environment: Challenges to Maritime Security,” the President of the Naval War College, Vice Admiral Arthur K. Cebrowski, offers several ways of conceptualizing the “domain” in which the U.S. Navy operates. “Of them, a less traditional, but compelling, approach is electronic. If one plots on a blank sheet of paper the location of radio emissions related to commerce, the result is a map of the world, with very prominent zones coinciding with conflict fault-lines. An analysis of the world’s finely tuned network of shipping produces a pattern complementary with the zones of most intense electronic activity. . . . All this produces a coherent pattern, one that suggests that the U.S. Navy will continue to be forward deployed and that gives a pretty good indication of where those deployments will be.” Our cover (adapted from the briefing graphic, produced by the College’s Graphics Department, to which the Admiral spoke) depicts that coherent pattern and, in so doing, points to this issue’s focal theme—forward deployment and the U.S. Navy. Dr. Daniel Gouré of the Lexington Institute challenges the necessity for forward deployment and the capacity of the Navy to maintain it. Prof. James F. Miskel, of the Naval War College faculty, takes a different view, accepting the requirement but arguing the need for a rigorous (ideally economic) metric if the U.S. Navy is to “be there” in the places in which it can best support the nation’s interests.

An early iteration of the Admiral’s presentation was reprinted as our Summer 1999 “President’s Notes.” A link to the graphics can be found on the College’s home page, http://www.nwc.navy.mil.
The Naval War College Review was established in 1948 as a forum for discussion of public policy matters of interest to the maritime services. The thoughts and opinions expressed in this publication are those of the authors and are not necessarily those of the U.S. government, the U.S. Navy Department, or the Naval War College.

The journal is published quarterly. Distribution is limited generally to commands and activities of the U.S. Navy, Marine Corps, and Coast Guard; regular and reserve officers of U.S. services; foreign officers and civilians having a present or previous affiliation with the Naval War College; selected U.S. government officials and agencies; and selected U.S. and international libraries, research centers, publications, and educational institutions.

Contributors

Please request the standard contributors’ guidance from the managing editor or access it on-line before submitting manuscripts. The Naval War College Review neither offers nor makes compensation for articles or book reviews, and it assumes no responsibility for the return of manuscripts, although every effort is made to return those not accepted. In submitting work, the sender warrants that it is original, that it is the sender’s property, and that neither it nor a similar work by the sender has been accepted or is under consideration elsewhere.

Permissions

Reproduction and reprinting are subject to the Copyright Act of 1976 and applicable treaties of the United States. To obtain permission to reproduce material bearing a copyright notice, or to reproduce any material for commercial purposes, contact the editor for each use. Material not bearing a copyright notice may be freely reproduced for academic or other noncommercial use; however, it is requested that the author and Naval War College Review be credited and that the editor be informed.

The Naval War College Review is listed in Ulrich’s International Periodicals Directory, Military Media Database, Free Magazines for Libraries, and by Oxburgh Communications; it is microformed by UMI® (Bell & Howell Information and Learning); it is abstracted in International Political Science Abstracts, in the Lancaster Index to Defence & International Security Literature, and in the Joint Warfighting Center Futures Database. It is indexed in the Air University Index to Military Periodicals, in Historical Abstracts and America: History and Life (both ABC-CLIO), the International Bibliography of Periodical Literature, the International Bibliography of Book Reviews, and selectively in the American Foreign Policy Index and INFO-SOUTH. A bound 1948–1991 index of Review feature articles is available from the editorial office; it is updated annually in the Winter issue. A 1948–2001 index is available on-line. Book reviews (beginning with Spring 1994) are indexed in the Gale Research, Inc., Book Review Index and on-line for recent years.


Periodicals postage paid at Newport, R.I. POSTMASTERS, send address changes to: Naval War College Review, Code 32S, Naval War College, 686 Cushing Rd., Newport, R.I. 02841-1207.

ISSN 0028-1484
CONTENTS

President’s Forum .......................................................... 7
More than ten years ago, historian Colin Gray asserted that, to put it bluntly, America has become accustomed to playing “away games.” America has enjoyed the great strategic advantage of deciding which wars to fight and whether to win or lose them. But now, many argue that it will not be so for long: the preferred American way of war is increasingly at risk.

Forward Presence and the Navy

The Tyranny of Forward Presence ................................. 11
Daniel Gouré
A cardinal principle of Navy strategic planning is that “shaping” the international environment is a necessary and appropriate mission. However, the ever-increasing scope of forward presence exerts a tyrannical hold on the future of the Navy, a hold that threatens—in an era of constrained defense budgets and rapidly changing threats—to break the force.

“Being There” Matters—But Where? ............................. 25
James F. Miskel
The driving factor in determining the requirements for routine, noncrisis forward presence ought to be reasoned, objective judgments about the relative importance of the various regions of the world to the United States. One approach is to define specific criteria—and for the time being, the most sensible criteria are economic.

India and Pakistan

Thinking about the Unthinkable ................................. 40
Paul D. Taylor
In 1998 and 1999 a series of simulations and “decision events” examined the possible consequences of recent ominous developments in South Asia. The most compelling result was that none of the experts, from a variety of backgrounds, nations, and organizations, argued that the scenario—a dispute over Kashmir leading to a nuclear exchange—could not happen.

“Security Is Like Oxygen”
A Regional Security Mechanism for West Africa ............. 52
Lieutenant Commander Seth Appiah-Mensah, Ghana Navy
West Africa envisages an ambitious collective-security mechanism for the subregion. However, the West African subregion may not yet have the political, military, or economic means to accomplish its strategic objectives. It will have great difficulty achieving them without increased assistance—properly coordinated and synchronized—from the United States, other states, and private institutions.
Future Carrier Aviation Options
A British Perspective. .......................... 63
David J. Jordan
Particularly in the 1964–70 period, the Royal Navy deemed the aircraft carrier an expensive irrelevance that could not be justified. Since then the aircraft carrier has returned to prominence, with the 1998 Strategic Defence Review (SDR) planning for two carriers. Nonetheless, a number of issues need to be settled before the new vessels enter service in 2012.

Transforming the U.S. Armed Forces
Rhetoric or Reality? .............................. 85
Thomas G. Mahnken
The U.S. armed forces themselves have embraced—at least rhetorically—the need to transform so as to meet the demands of information-age warfare. Nevertheless, in each service, efforts at transformation have faced opposition from traditionalists who perceive threats in new ways of war.

New Insights from Old Books
The Case of Alfred Thayer Mahan .......................... 100
Jon Sumida
The major arguments of Mahan are very different than has been supposed. Yet the issues that prompted him to put pen to paper remarkably resemble those of today—joint operations in coastal waters, global free trade, uncertainty about what America’s proper naval role should be, and a “revolution in naval affairs.”

The Coast Guard and Navy
It’s Time for a “National Fleet” .......................... 112
Colin S. Gray
The U.S. Navy and the Coast Guard are natural complements, and this complementarity should be expressed practically, in a “national fleet” that is not mere rhetoric. The argument is new, but it is well established in historical practice. What should the Navy and Coast Guard do to aid each other?

Set and Drift
Doctrine Matters
Why the Japanese Lost at Midway .......................... 139
Jonathan B. Parshall, David D. Dickson, and Anthony P. Tully
The Perils of Paperless .............................. 152
Some Questions about the Latest Defense Business Trend
Patrick J. Geary

https://digital-commons.usnwc.edu/nwc-review/vol54/iss3/30
In My View ................................................................. 157

Review Essay
Where Will Presidential Autocracy Take Russia? ............................. 165
The Russian Presidency: Society and Politics in the Second Russian Republic,
by Thomas Nichols
reviewed by Sergei Khrushchev

Book Reviews
The Coming Crisis: Nuclear Proliferation, U.S. Interests, and World Order,
edited by Victor A. Utgoff
reviewed by Jon Greene .................................................. 169
Planning the Unthinkable: How New Powers Will Use Nuclear, Biological,
and Chemical Weapons, edited by Peter Lavoy, Scott Sagan, and James Wirtz
reviewed by K. A. Beyoghlow .............................................. 170
Chemical-Biological Defense: U.S. Military Policies and
Decisions in the Gulf War, by Albert J. Mauroni
reviewed by Pietro D. Marghella .......................................... 172
The Changing Role of Information in Warfare,
edited by Zalmay M. Khalilzad and John P. White
reviewed by Eric J. Dahl .................................................... 174
Identifying Potential Ethnic Conflict, by Thomas S. Szayna
reviewed by Paulettta Otis .................................................. 176
Interpreting China’s Grand Strategy: Past, Present, and Future,
by Michael D. Swaine and Ashley J. Tellis
reviewed by Carmel Davis .................................................. 178
reviewed by Robert Morabito .............................................. 179
Flying Black Ponies, by Kit Lavell
reviewed by William M. Calhoun .......................................... 181
How Effective Is Strategic Bombing?
Lessons Learned from World War II to Kosovo, by Gian P. Gentile
reviewed by Phillip S. Meilinger .......................................... 182
Norstad: Cold War NATO Supreme Commander—
Airmen, Strategist, Diplomat, by Robert S. Jordan
reviewed by Douglas Kinnard ........................................... 183
American Airpower Strategy in Korea, 1950–1953, by Conrad C. Crane
reviewed by Wilfred F. Brown ........................................... 185
Decoding History: The Battle of the Atlantic and Ultra, by W. J. R. Gardner
reviewed by Christopher Bell ............................................................... 186

Dog Boats at War: Royal Navy D Class MTBs and MGBs, 1939–1945,
by Leonard C. Reynolds
reviewed by Carl Otis Schuster .......................................................... 187

The Legacy of the White Oak Laboratory, by William B. Anspacher et al.
reviewed by Frank C. Mahncke ......................................................... 189

From the Editors .............................................................................. 192

Of Special Interest .......................................................................... 194
Vice Admiral Cebrowski commanded Fighter Squadron 41 and Carrier Air Wing 8, both embarked in USS Nimitz (CVN 68). He later commanded the assault ship USS Guam (LPH 9) and, during Operations DESERT SHIELD and DESERT STORM, the aircraft carrier USS Midway (CV 41). Following promotion to flag rank he became Commander, Carrier Group 6 and Commander, USS America Battle Group. In addition to combat deployments to Vietnam and the Persian Gulf, he has deployed in support of United Nations operations in Iraq, Somalia, and Bosnia. He has served with the U.S. Air Force; the staff of Commander in Chief, Atlantic Fleet; the staff of the Chief of Naval Operations, on four occasions; with the Joint Staff (as J6); and as Director, Navy Space, Information Warfare, and Command and Control (N6). Vice Admiral Cebrowski became the forty-seventh President of the Naval War College in July 1998.
America's international status and global interests require that our forces have as their operating domain not only the majority of the earth's surface, but also the skies and space above, the ocean depths below, and the electronic environment we think of as cyberspace.

The title of these notes might well be “Stronghold Lost.”

During our March intersessional conference, the importance of geography in the Information Age was debated. The U.S. military has already demonstrated that, as information is substituted for mass, military forces can respond more rapidly and overcome many of the traditional impediments imposed by geography. We further hear that in Information Age war the occupation of terrain may not always be a legitimate objective of military operations. These arguments combine to challenge traditional thinking about the priority of geography—strategically, operationally, and tactically. Care must be taken lest we make judgments which are too broad, resulting in decisions which increase risk.

The concept of a stronghold is closely linked to key elements of geography in both time and space. Strongholds are places of security for friendly forces, areas that the enemy cannot reach and his forces cannot effectively threaten. This concept is deeply imbedded in military thinking. Traditionally, strongholds were situated where geographic features suggested, as at West Point on the Hudson River during the American Revolution or Fort McHenry protecting Baltimore’s Inner Harbor in 1814; sometimes, however, they were created as military expedience dictated, as in “circling the wagons” on the American Great Plains. Either way, to be able to fight from a stronghold is a great advantage. To allow an enemy a stronghold is to yield to him some degree of control over the initial conditions of an engagement. Because of the way that a war proceeds, small changes in initial conditions can result in profound changes in outcome. This is why controlling
the initial conditions, holding the initiative, and fighting from strongholds are so important.

More than ten years ago, historian Colin Gray asserted that America has had the wonderful blessing of never having to fight for its survival at the start of a war. To put it bluntly, America has become accustomed to playing “away games.” Bordered by two vast oceans and two friendly neighbors, America has enjoyed the great strategic advantage of being a secure stronghold. With that advantage, America has decided which wars to fight and whether to win or lose them. But now, many argue that even if Gray still is right today, he will not be so for long: the preferred American way of war is increasingly at risk.

The prospect of national vulnerability affects the security policies and decisions of the new administration. It casts a different light on the controversial subject of missile defense and places it in the context of a far larger issue. The loss of the homeland stronghold indicates a requirement for a broad program, of which missile defense is but a part. A global power with global interests quite naturally wants to be able to use all the instruments of national power for policy purposes ranging from homeland defense through the projection of national values. But, in the case of America, the growing vulnerability of our nation itself may restrict, if not exclude, some options at all levels of security planning.

Our national preference (and a moral imperative) is to deliberate at length whether to go to war. Once war is decided upon, this nation traditionally builds a coalition, preferably including nations near the scene of action, and moves forces forward en masse. A forward stronghold is created with the acquiescence and assistance of coalition partners in the region. Alternatively, a stronghold can be created at sea. Commanders maneuver their forces to create secure operational zones if only for brief, necessary periods.

Thus, we see the concept of stronghold manifested at the strategic, operational, and tactical levels of war. Americans prefer this approach for the same reason that potential enemies seek to prevent us from using it. In fact, we like the concept of strongholds so much that the nation is willing to go to extraordinary lengths to create and maintain them. I hardly need to remind readers that in World War II we suspended the civil rights of Japanese-Americans because their very presence was thought to be an internal threat to the national stronghold. During the Cold War, we adopted the seemingly perverse strategy of mutually assured destruction in the hope that, with both sides’ strongholds vulnerable to devastation, neither side would dare attack at that level.

Today, for example, at the highest governmental level the use of chemical or biological weapons is declared to be unacceptable, and it is a national goal that certain technologies will not be allowed to proliferate. By invoking prohibitions in this way the nation seeks to establish universally acknowledged strongholds.
A strategic example is the decision to dominate the sea, air, space, and cyberspace as operating domains. At the level of grand strategy we speak of the concept of “projecting defense” as a means of coalition creation and assurance. This defensive concept extends to the operational level of war when we employ forward basing, either ashore or afloat. At the tactical level of war, strongholds can be created in locale with maneuver, or in time with speed. The techniques used normally depend on the physics of the systems involved and on the environment. Tactically, the duration of a stronghold may need to be only seconds, or perhaps less.

What are the implications for the Navy? From the perspective of strategy, the Navy must train, equip, and operate to maximize its capabilities in its operating domain. America’s international status and global interests require that our forces have as their operating domain not only the majority of the earth’s surface, but also the skies and space above, the ocean depths below, and the electronic environment we think of as cyberspace. Here, the U.S. Navy must predominate.

At the operational level of war, the Navy must guarantee both speed of deployment and speed of employment of force. Time lost in the creation of a stronghold in effect grants a stronghold to the enemy. This suggests a need for forces which are capable of clearing distant seas of mines and enemy submarines, while projecting air and missile defense and essential elements of the joint force, including sensors and command and control capabilities, all at high speed in the opening days or even hours of a conflict.

The critical planning requirement at the tactical level of war is to offset the emerging condition of tactical instability. Tactical instability occurs when unit or force capability is allowed to increase disproportionately to survivability, which is staying power under stress. What tactical instability produces is a force which has to be risk averse. Such a force is unable to conduct sustained operations in heavily contested areas; by extension, then, it is unable to support other elements of the joint force in such areas. In the case of the U.S. Navy, if highly capable units are too vulnerable to operate in littoral regions, the Navy cannot perform missions itself nor support joint and coalition forces in those regions. The point is not that the Navy must be able to create a permanent geographic stronghold in the close-in littoral for any and every purpose, but that it must be able to fight and sustain forces whenever and wherever that is required. The consequence of tactical instability will be that an enemy will enjoy the advantages of operating from a stronghold, one that contains the vast majority of the world’s population, the sources of much of the world’s wealth, and the termini of the world’s most critical networks.

We—not our enemies—must fight from the strongholds of the future. To do so will require a profound comprehension of tomorrow’s battlespace, from the complexities of urban combat to the intricacies of orbital dynamics, and from
the details of electromagnetic propagation to the subtleties of an adversary's motivations. Will we be able to establish strongholds wherever and whenever the United States may need them? That is the challenge now being posed to tomorrow's leaders.

ARTHUR K. CEBROWSKI

Vice Admiral, U.S. Navy
President, Naval War College
A specter is haunting U.S. Navy strategic and force planning. It is the specter of forward presence, the continual deployment of Navy and Marine Corps units in waters adjacent to foreign littorals. Although the Navy speaks of its central purpose as maritime power projection, it is forward presence, particularly in peacetime, that drives both force structure requirements and operations tempo. The demands placed on both force structure and operations tempo by the Navy’s long-standing commitment to maintain forward presence in multiple regions have been exacerbated in the past few years by that institution’s desire to extend its area of influence to both littoral waters and the land beyond. The ever-increasing scope of forward presence exerts a tyrannical hold on the future of the Navy, a hold that threatens—in an era of constrained defense budgets and rapidly changing threats—to break the force.

The general argument for forward presence as a cardinal principle of Navy strategic planning is that “shaping” the international environment is a necessary and appropriate mission for the U.S. military in general, and the Navy in particular.¹ The military is not alone in believing in the importance of the “shaping” mission. Under various rubrics, this impetus was central to the Clinton administration’s articulation of national security policy and national military strategy.² Were this only the view of one administration, it could be readily dismissed as international social work.³ But a growing chorus of voices in the academic and analytic communities argues...
that U.S. defense planning should emphasize “shaping” functions. Some are so bold as to speak of a new role for U.S. forces in terms of “what can only be termed ‘imperial policing.’”

The myth that the world is in dire need of shaping or policing derives from the proposition that with the end of the Cold War the forces that had dampened disorder and disunity ceased to function. This “chaos theory” increasingly pervades all the services and the Department of Defense as well, but the Navy and Marine Corps have been among its chief proponents. Here is but one example of the Navy–Marine Corps view:

Never again will the United States exist in a bipolar world whose nuclear shadow suppressed nationalism and ethnic tensions. The international system, in some respects, reverted to the world our ancestors knew. A world of disorder. Somalia, Bosnia, Liberia, Haiti, Rwanda, Iraq, and the Taiwan Straits are examples of continuing crises we now face. Some might call this period an age of chaos.

But is this Hobbesian vision real? Has the world reverted in the last decade to a state of nature, from some prior regime of civility, or at least restraint? The Middle East suffered four Arab-Israeli wars prior to the end of the Cold War. For decades, Iraq engaged in predatory behavior toward its neighbors—producing most notably a ten-year bloodbath with Iran—before deciding to invade Kuwait. India and Pakistan have several wars to their account, the last in 1971, as well as chronic conflict over Kashmir. The Taiwan Straits is a military problem not because of the end of the Cold War but because of China’s arms buildup and the failure of the United States to provide countervailing capabilities to Taiwan. The civil and regional wars of Africa are largely consequences of colonization and the rivalries of the Cold War itself.

Many once-fractious parts of the world have become more stable over the past decade. The collapse of the Warsaw Pact and the downfall of the Soviet Union eliminated the major supporter of international terrorism. Thereafter, the inability of Russia to provide cheap conventional weapons to client states also reduced regional arms races dramatically. Lack of arms may have reduced as well the aggressive tendency of such former client states as Syria, Libya, Iraq, and Iran. One can point even to recent events on the Korean Peninsula as a direct, albeit delayed, result of Pyongyang’s loss of its Soviet godfather.

Where problems have arisen, it is not clear that the end of the Cold War was the catalytic event. It is difficult, for instance, to establish a correlation between the end of the Cold War and the rise of militant Islam. Further, events in Indonesia have had less to do with the rise and fall of superpowers than with the consequences of the Asian economic crisis (during which, it must be noted, the
Treasury Department did more to maintain stability than all the U.S. forces deployed to the region).

Current military planning has somewhat tempered its earlier “Boschian” vision of global chaos, asserting now that it is the uncertainty of our time and the difficulty of predicting the future security environment that necessitates a strategy of power projection based on forward presence. The fault, in that view, lies not in the unstable nature of the external world but in our inability to forecast the future accurately. For planning purposes, uncertainty may be as good as chaos. In some respects it is even better, since—as the services’ planning documents note—it requires that the military maintain capabilities to address all threats.

This sense of chaos, or even mere uncertainty, masks what is really happening: a restructuring of the international environment, the creation of a new international system. We know from history that such restructuring is long, complex, and often quite messy. Wherever we look, in each of the critical regions of the world, the character of the relations among the dominant powers has yet to be firmly set, much less put on a course toward stable, positive, and peaceful relations. Western Europe is waiting to see if a closer union, and with it an incipient common security and defense identity, can be effected. Nato expansion is confronting the question of Russia’s legitimate security interests in Eastern Europe. China’s role in East Asia is being defined by Beijing—witness the 1999 military maneuvers and missile launches against Taiwan—in ways that must make all of its neighbors nervous; how China acts will determine to a large extent the behavior of others in the region. The relationship between India and Pakistan is as tense as it has ever been; increasingly, both states see the need to reach out to other powers of the Middle East and Asia in order to strengthen their positions in their own rivalry. Finally, the future of the political and security relationships in the Persian Gulf is frozen, and it will be as long as Iraq and Iran remain pariah states and the United States is required to maintain a military presence in the area.

Historically, the creation of new international orders has been dominated by major economic and military powers. This current period of evolution appears to be no different. In prior periods of reorganization, emerging powers have sought ways to shift regional and even global power balances in their favor, provoking similar behavior by their adversaries. (It is in this light that we need to look with some concern at current Russian and Chinese efforts to forge a new strategic alliance.)
Certain regions will be most important in the development of the new international order. For future U.S. policy, three regions are of vital importance: Europe, from the Atlantic to Russia’s borders; the Pacific Rim, from Korea through Southeast Asia; and South Asia and the Persian Gulf. Those regions have three things that set them apart from the rest of the world. First, they contain the overwhelming predominance of global wealth, economic activity, and technological investments. Second, they are the loci of vital U.S. allies and of economic interests that must be protected. Third, they each border on one or more of the emerging potential competitor states.

The United States is the sole global power; it has interests in every region of the world and vital interests in each of them. Thus, while it is difficult to identify where confrontations will arise, the sheer breadth and scope of U.S. interests abroad provide more than a few reasons that this nation may find itself at basic odds with local adversaries. Indeed, at least one major study of U.S. foreign policy in the next century argues that the foremost U.S. interest in Asia and Europe is to prevent the domination of those regions by adversarial powers. Therefore, the United States could find itself in confrontations with rising powers as it seeks to preserve regional balances of power or American access. This would be particularly likely should, as has been the case in the past, a powerful regional state threaten U.S. allies. The United States is likely to be the only nation that can provide sufficient military support to enable these allies to deter or, if necessary, defeat such an adversary.

It is true that the post–Cold War world has demonstrated a degree of disorderliness. But it can hardly be said that the world has entered a period of mounting chaos. Nor can it be claimed that U.S. decision makers and planners are paralyzed by uncertainty. They continue to make decisions and set priorities on force structure, regional deployments, and future acquisitions with a great deal of self-assurance. The chaos/uncertainty argument, then, serves largely as a means of defending the military against the increasingly evident need to make hard choices with respect to current missions and future capabilities. For the Navy, the validity of the doctrine of forward presence represents one of those hard choices.

SHOULD THE NAVY MAINTAIN A STRATEGY OF FORWARD PRESENCE?

It is not clear that the U.S. military should focus its planning and force-building around forward presence, much less “imperial policing.” The idea that military forces can shape the political environment in regions in which they are deployed has become fashionable as a result of the rise of an issues-based approach to national security policy. Many of these issues are sociopolitical in nature, and
their solutions fall, broadly speaking, under the heading of “shaping.” The trend toward employing military forces for political purposes has been given additional impetus by the activism of the regional commanders in chief (such as those of Pacific Command or Central Command), which has grown as the power of the State Department and U.S. ambassadors to conduct foreign policy has declined.11 (One of the potential consequences of their use of forward-deployed forces for political purposes was highlighted by the USS Cole incident.)

It is for these reasons, then, that the U.S. military is increasingly focused on and driven by the demands of peacetime and crisis forward presence. The problem of maintaining forward presence has been a crucial factor, for instance, in the U.S. Air Force’s creation of a new organization centered on ten aerospace expeditionary forces. The U.S. Army is undergoing its own transformation, seeking to become more responsive and deployable. Each of the services is investing in capabilities to make rapid forward presence easier to establish, whether for major conflicts or smaller contingencies.

In particular, and without question, forward presence has served the Navy well. Forward presence provides a defensible rationale for force sizing, a matter of particular importance in the absence of a threat.12 In any case, the Navy functions best when it is under way, and as long as it is steaming, it might as well do so where it might be needed.

The idea of forward presence, however, is for the Navy more than a bureaucratic convenience; it is an article of faith. According to the Navy’s own Strategic Planning Guidance, “By remaining forward, combat-credible naval expeditionary forces guarantee that the landward reach of U.S. influence is present to favorably shape the international environment.” In the Navy’s view, forward-deployed naval forces discourage challenges to U.S. interests, deter would-be aggressors, and, should deterrence fail, provide means for a timely response. For these reasons, the Navy argues, it could play a new and unique role in U.S. national security. But for this to be true, forward presence has to be the Navy’s central mission.13

For a number of reasons, tying the future of the Navy to forward presence is problematic. The concept of “shaping” the international environment is fuzzy at best. Too often it has extended well beyond traditional notions of security to involve, inter alia, attempts to influence the internal politics of failing states, efforts to address almost intractable socioeconomic problems, and engagement in what are classic policing functions. Looked at this way, Navy combat forces seem to have little relevance.14 The forces that would seem to be most useful in the social-work and policing dimensions of forward presence are those generally classed as “combat support” or “combat service support” (e.g., engineer, military police, logistical, and medical units).
The term “forward presence” too is subject to interpretation and competing definitions. In its narrow sense, the emphasis is on forward—it simply means the deployment of forces in proximity to locations of interest to U.S. security and foreign policy. A broader definition, focusing on the word presence, suggests more complex and political purposes, for which presence generally needs to be nearly continuous and highly visible—requirements that can limit both the flexibility and the combat effectiveness of the forces engaged.

Leaving aside for the moment the question of what constitutes a combat-credible force, it is fair to ask what evidence there is that naval forward presence helps to shape the international environment. One can acknowledge that military forces can perform tasks that are essentially political in nature, such as demonstrating resolve and commitment. The objective of these tasks is different from that of forward presence, as narrowly defined above.

Advocates of forward presence as an instrument of U.S. foreign policy must acknowledge that there is no empirical evidence to support their case. This is particularly true for naval forward presence. While various theories have been propounded as to the relationship between the pursuit of national objectives, the protection of regional interests, the suppression of sources of regional instability, and forward presence, none has any real data to support it. It has been possible to show in certain instances some relationship between the ebb and flow of economic indicators and the deployment of U.S. forces; however, these cases involve the deployment of forces after crises or conflicts have started. Such analyses have not been able to demonstrate the usefulness of peacetime forward presence as a mechanism for preventing conflicts and shaping regional environments. As one analyst (in fact, an advocate of naval power) noted a few years ago, “The interesting fact is that there is virtually little or no evidence, analysis and rigorous examination on which to make a fair and objective assessment of the benefits, costs, advantages and downsides of presence... [T]he record is at best ambiguous regarding the utility, benefits and disadvantages of naval presence.”

Even the projection of maritime power may not serve to shape the environment or resolve a regional crisis. The history of the U.S. presence in the Persian Gulf in the 1980s—including Operation EARNEST WILL, the ill-fated attempt to protect oil shipments by reflagging foreign-registry tankers—does not support the thesis that naval forward presence exercises a positive influence on regional dynamics. Similarly, it is considered self-evident in Navy circles that the
deployment of two aircraft carriers to the Taiwan Straits region ended the 1996 crisis. At least one post-incident assessment suggests otherwise.\footnote{18}

In addition to the shaping function, the Navy asserts, forward presence provides unique operational advantages. The Navy makes a strong case that such deployments are critical enablers of joint warfare, through a combination of sea control and maritime power projection; for instance, where land bases are not available, naval forces can become alternative bases. Naval power-projection capabilities, in this view, are likely to be less vulnerable to adversary attack than land bases. Even here, however, the other services have attempted to make cases that forward presence can be accomplished in other ways and with different means.\footnote{19}

The land-versus-sea-base argument has been going on for a long time, with no resolution in sight. It is sufficient here to point out that the fact that naval forward presence may be needed if land bases are not available does not make it the preferred solution. Indeed, when the stakes are sufficiently grave or vital interests and allies are threatened, it is unlikely that U.S. political and military leaders will rely solely on naval forward presence. To put it bluntly, if land bases are necessary, they will be found or even seized. This is an often-overlooked lesson of the Gulf War and the Kosovo campaign. In this connection, the Navy itself speaks of its role as that of an enabler, suggesting that it is the responsibility of the other services—those that require land basing—to win a war. In that light, it is not clear that allies will find the simple presence of naval units offshore adequate. U.S. “boots on the ground” have reassured allies for some fifty years as indications that the United States is willing to share equally in the risks of resisting aggression.

At the very minimum, the Navy needs to rethink how it describes the forward-presence mission.\footnote{20} Justifying forward presence in terms of the ability to shape the international environment raises questions of how relevant the current Navy force structure is to that purpose. Moreover, it risks promising more than the Navy can deliver, at least in terms of demonstrable impact. Also, because forward presence is tied to a particular national security strategy, it may be rendered less relevant if the new administration formulates a new, more restrained strategy.

It is, then, difficult to see continuous, peacetime forward presence as anything other than a vehicle for defending the Navy’s desired force structure. The political rationale is weak at best, and holding on to it may undermine the Navy’s case for more capable forces in the future. One naval officer appears to have recognized the danger in a recent article: “If . . . naval forward presence forces have but small roles in crisis response and contingencies, such forces are luxuries that may have
some relevance in peacetime diplomacy but little usefulness in crisis and war. This is not an impression that bodes well for the future of a military service."

**CAN THE NAVY MAINTAIN A STRATEGY OF FORWARD PRESENCE?**

Even if it were obvious that forward presence is an important tool of U.S. national security strategy, there are reasons to believe that it will not be possible to continue it for long. Forward presence places inordinate and, in the current budgetary environment, unsustainable physical demands on the Navy. Some fixed and substantial number of ships is necessary to maintain a fraction of them on station continually. For every ship deployed, the U.S. Navy requires between three and five more in rotation: steaming to or from the deployment area; in overhaul; in port for leave and repair; and “working up” in local training exercises. All that in turn translates into a minimum required budget. It is clear that the Navy will not have a large enough budget, and thus not enough ships. Vice Admiral Edmund Giambastiani was reported to have pegged the Navy–Marine Corps annual procurement budget at between twenty-eight and thirty-four billion dollars annually, far above the twenty-two-billion average for the past decade. The lower procurement number translates into reduced ship construction and, inevitably, a navy of fewer than three hundred ships. Even if additional funds and an adequate number of ships were available, changes to the threat environments in regions where forward naval presence is now practiced raise questions as to its wisdom.

All naval forces are subject to the terrible tyranny of distance. It takes time for ships to sail from their home ports to deployment areas. Nowhere are the distances to be traveled greater than in the Pacific. Whereas it typically takes a U.S. warship about eleven days to travel from the East Coast to its assigned station in the Mediterranean, the same deployment can take up to twenty days from the West Coast of the United States to the littoral waters of the Asian landmass.

No other navy is so tyrannized by its strategy and geography as that of the United States. Every other naval power is concerned largely with the protection of its own coastlines and nearby waters. Only the United States is confronted with the need to project naval power eight to ten thousand miles to areas of concern. The farther away a deployment area is from home ports, the more ships are required in order that a given number can be continually present. Hence a strategy that emphasizes forward presence inevitably puts additional strain on an already-overstretched U.S. Navy.

From a force of nearly six hundred ships in the late 1980s, the Navy has been reduced to a little over three hundred ships today, of which approximately 45 percent must be under way in order to meet current peacetime responsibilities.
This places enormous strain not only on the ships but on the men and women who serve aboard them. At the same time, because of reduced funds for shipbuilding, the average age of the Navy’s vessels is increasing; accordingly, breakdowns become more frequent, maintenance costs rise, and availability rates decline. However valuable forward presence may be in the Pentagon’s internecine budget battles, it can impose intolerable stress on a service that is asked to perform missions for which it is underequipped. When forward presence becomes a burden to the very service that is its chief proponent, it is time to rethink the whole proposition.

The Navy understands the problem. In testimony before the House of Representatives in 2000, Vice Admiral Conrad Lautenbacher, Deputy Chief of Naval Operations, declared that “it is no secret that our current resources of 316 ships are fully deployed and in many cases stretched thin to meet the growing national security demands.”

This is not merely the view from headquarters. Admiral Dennis McGinn, commander of the Third Fleet, stated before Congress in February 2000 that “force structure throughout the Navy is such that an increased commitment anywhere necessitates reduction of operations somewhere else, or a quality of life impact due to increased operating tempo.” The commander of the U.S. Fifth Fleet, operating in the Arabian Sea and Persian Gulf, said it best:

Although I am receiving the necessary forces to meet Fifth Fleet obligations, the fleet is stretched and I am uncertain how much longer the Navy can continue to juggle forces to meet the varied regional requirements, including Fifth Fleet’s. I am uncertain that we have the surge capability to meet a major theater contingency, or theater war. Eventually, the increased operational tempo on our fewer and fewer ships will take its toll on their availability and readiness.

The reality is that numbers matter. The U.S. Navy is critically short of ships; it does not have enough to maintain a full-time, combat-credible naval presence in regions of interest to the U.S. and provide the necessary surge capability for crisis or war. As a result of recent events like Kosovo, for which the western Pacific was stripped of its aircraft carrier, public and congressional attention has been focused on the inadequacy of the Navy’s inventory of carriers. Further, the Joint Chiefs of Staff have published a study concluding that the nation requires sixty-eight attack submarines instead of the fifty that have been allowed. A recent surface combatant study concludes that the Navy requires up to 139 multimission warships in order to satisfy the full range of requirements and carry out day-to-day operations; instead, the Navy has been allowed only 116.
At least a quarter of its surface combatants are aging frigates and older destroyers that lack offensive and defensive capabilities essential to a twenty-first-century navy. Speaking of the lack of surface combatants, one senior naval officer has been quoted as saying, “We know we are broken. We are running our ships into the ground, our missions are expanding and our force structure is being driven down to 116 surface ships. We have to address it before we hit the precipice.”

Unfortunately, without significantly higher defense budgets, there is no possibility that the Navy will be able to acquire the ships and submarines it needs to maintain its current forward presence posture. It is already evident that U.S. defense spending is well short of what will be required to maintain the existing force structure. The United States must be willing to spend on average 4 percent of its gross domestic product (GDP) to support fully the force recommended by the Quadrennial Defense Review over the next twenty years, fiscal years (FY) 2001–20. In fact, however, based on the current FY 2002 budget submission to Congress, defense spending will fall from 2.9 percent of GDP in FY 2000 to 2.4 percent in FY 2010, and to 2 percent in 2020.

The Congressional Budget Office reports that the Defense Department is faced with annual budget shortfalls of fifty-two to seventy-seven billion dollars. General Henry Shelton, chairman of the Joint Chiefs of Staff, testified before Congress in October 2000 that the military services had estimated that they will need at least $48.5 billion more each year. The Secretary of the Air Force, F. Whitten Peters, asserted in a recent interview that the U.S. military needed some $100 billion over current spending levels in order to replace aging equipment and maintain or improve operational readiness. Unless real annual defense spending is increased well above the current $310 billion at some time during this decade, the president and Congress will be left with little choice but to make additional personnel cuts, force structure reductions, and base closures.

The Navy will suffer severely if such projections, and others, of budgetary shortfalls are even approximately accurate. A recent Navy study warned that procurement was short some eight-five billion dollars for the period 2008–20, with the shipbuilding budget likely to be underfunded by some four billion annually, and naval aviation by $3.3 billion. These shortfalls could result in a Navy one-third to one-half its present size by the year 2010.

If the force cannot be recapitalized, perhaps it can be modernized or transformed, thereby avoiding the problem of finding the necessary additional funds. A number of analytic and political writers have advocated “skipping generations” in procurement in order to focus attention and resources on revolutionary...
capabilities. Unhappily, the idea of skipping a generation is a fantasy. There is an illusion among its advocates that the current force will last the additional twenty-odd years while the transformation takes place. In fact, however, the funds necessary to support a transformation can be freed up only if current forces and near-term acquisitions are sharply reduced. Reducing forces and acquisitions now will only make the conduct of current operations, including forward presence, more difficult. Moreover, reducing the acquisitions will seriously damage the defense industrial base, on which the services will have to rely for the production of next-generation equipment.

Budgetary strictures also constrain the fielding of the advanced capabilities forward-deployed forces will need if they are to be combat credible and survivable. The Navy acknowledges that the threat to its forward-deployed forces is serious and likely to grow substantially worse over the next few decades. This means that combatants built for the Cold War are increasingly vulnerable, particularly in littoral waters. The Navy will need to invest in a host of new technologies enhancing both the offensive and defensive power of the fleet; otherwise, forward presence will be not merely an expensive conceit but a truly dangerous fetish. Yet it is not clear that either the technology or the resources will be available. The demand that the Navy operate forward in peacetime, then, exerts a perverse effect, forcing on the Navy an expensive modernization/transformation effort that may in the end prove unsuccessful, if only due to a lack of funds.

It must also be recognized that even if transformation is possible, it will take decades to complete. As a result, today’s Navy will be required to execute the forward-presence strategy ten and even twenty years into the future. If, as is argued by advocates of transformation, today’s Navy will be the wrong force with which to maintain forward presence or contest littoral waters, it seems obvious that the problem is not with the force but with the demand that the Navy continue to base its strategy on forward presence. The Navy must seek ways other than slavish obedience to the tyranny of forward presence to pursue its strategic objectives and support national security.

There remains a final question. Facing a growing littoral threat, depending on large “Cold War era” ships and submarines, and recognizing the effort by some potential adversaries to acquire “green” and even “blue-water” capabilities, why does the Navy continue to emphasize forward presence? It would seem reckless, to say the least, to continue to pursue a demanding strategy with declining resources of the wrong type. Moreover, it would seem to be a waste of the single advantage that the U.S. Navy possesses and that will remain uncontested for decades to come: its ability to dominate the open oceans.

Operating in close-in waters would appear to provide littoral adversaries with an unacceptable advantage. The desire of potential adversaries to contest
the U.S. Navy for control of these waters suggests that it would be foolhardy for the Navy to sail into that trap.

THE FUTURE OF FORWARD PRESENCE

The future of forward presence, then, appears uncertain at best. The American people’s patience with the idea that the United States can shape an international environment to suit its sensibilities appears to be wearing thin. A more judicious approach to the application of military power in the service of foreign policy will inevitably lead to a reduced requirement for forward presence. Where peacetime forward presence is required, naval forces may not be able to provide it more effectively than other kinds of forces. It is possible that policy makers and the public alike will look for more “bang for their presence buck.”

The Navy acknowledges that if forward-deployed forces are to play useful roles in peacetime or crisis, they must possess credible combat power. It is not clear how this can be accomplished in the face of the emerging threat. The proliferation of asymmetric and anti-access capabilities may threaten the survivability of forward-deployed naval forces. This problem is particularly acute for traditional surface platforms. Efforts to address the emerging vulnerability of forward-deployed naval forces by changing the character of naval systems and developing new concepts of operations may compromise the combat capability of such forces. To the extent that enhanced survivability must be acquired at the expense of offensive capabilities, it would seem to undercut the basic rationale for forward presence.

Finally, if forward-deployed capabilities can be maintained only at the expense of the ability to control the broad oceans, it will have proven to be a bad decision. At present there are no threats to the U.S. Navy in the open oceans, and this will be the case for the next several decades. However, a force built over the next ten or twenty years for forward presence and littoral combat will have to meet whatever threats emerge in the “shallow seas” for many decades beyond. Increased competition between the United States and rising regional powers could result in a challenge to the U.S. Navy’s mastery of the open oceans, or at least one ocean. Such a challenge could come soon enough to necessitate reconsideration of the present policy of optimizing naval forces for the forward-presence mission.

NOTES


20. Ibid., pp. 31–2.


Which parts of the world will be most important to the United States in the future? Where should the United States be ready to fight a war? During the Cold War the answers to these questions seemed obvious. The parts of the world that were most strategically important to the United States were the lands along the Nato–Warsaw Pact fissure in Central Europe, and Northeast Asia, where two allies, Japan and South Korea, abutted the two largest communist powers, the Soviet Union and the People’s Republic of China. Other parts of the world could become strategically important if events there involved a Soviet or Chinese client, or even a potential client. When this happened, that part of the world became, ipso facto, an area of great importance to the United States, leading sometimes to covert conflict and sometimes to large-scale fighting.

Today, if only because there no longer is a superpower for the United States to balance against, U.S. interest in some regions has diminished. This has, for example, been true with respect to Africa. While the United States and the Soviet Union played strategic chess against one another in the Third World, African countries were perceived as important. But once the chess game ended, the great powers lost interest in the rooks and pawns. Developments in Africa ceased having balance-of-power ramifications, so the United States and most other major powers started to pay less attention to them. Some other parts of the world, of course, remain of great interest to the United States, but now for economic reasons as much as, if not more...
than, traditional political-military balance-of-power reasons. One particular aspect of national security strategy—military forward presence—deserves reassessment in light of the shift that is taking place in security interests.

The driving factor in determining the requirements for routine, noncrisis forward presence ought to be reasoned, objective judgments by the president and Congress about the relative importance of the various regions of the world to the United States. Indeed, the challenge for national strategists is to determine whether the regions that are important to the nation today will be equally important tomorrow and, if not, to begin making the necessary adjustments in forward presence strategy. Meeting this challenge requires clear and objective thinking about where national interests may lie in the future. It requires thinking that is unencumbered by the traditional Euro-centricity of the American foreign policy establishment and media, or by the habit acquired in the Cold War of instinctively assuming that the national-interest glass in some regions is perpetually more than half full.

One way of forcing policy makers to take a more objective approach and limit the influence of nostalgic assumptions about military forward presence is to define specific criteria for determining the intensity of U.S. national interests. Until such time as the global balance of military power shifts and a serious global threat begins to emerge, the most sensible criteria are economic. For the purposes of reassessing American military forward presence strategies, the most telling statistics involve interdependence. The practical effect of adopting economic criteria would be to accord a higher priority to those parts of the world where economic interdependence with the United States is growing and a lower priority where it is diminishing.

Forward presence entails the deployment of military forces overseas on a full-time basis or on rotational cycles during noncrisis periods. When crises do occur in important regions, adjustments are often made in the pattern of forward presence deployments to signal high levels of concern or to deter military action by third parties. For example, in 1997 and 1998 military forces were deployed to Southwest Asia to increase the existing levels of forward presence in response to threatening gestures by Iraq against Kuwait. Another example occurred in 1996, when the United States deployed warships to the seas around Taiwan in response to escalated tensions between Beijing and Taipei. However, the principal focus of this article is noncrisis, routine forward presence—although crisis deployments such as these will be discussed.

Forward presence has two broad goals. The first is to help ensure regional stability and promote productive relations with the United States by demonstrating its national interest in a region or an individual state as well as its commitment to
the defense of friends and allies. From this perspective, both routine forward presence during noncrisis periods and special deployments in response to crises are as much diplomatic-political functions as they are military. To maximize the yield from this perspective, forward presence should be allocated primarily to those parts of the world that will be most important to the United States in the future. That is to say, the United States government should do what any sagacious investor would do: it should invest more heavily in areas where the prospective returns are great and only sparingly, if at all, where the potential for profit is slight.

The second broad goal of both routine and crisis-oriented forward presence is to improve the ability of U.S. military forces to fight and win wars that are not deterred. Forward presence does this by providing opportunities to promote interoperability between U.S. and friendly forces. Forward presence also increases the operational awareness by American military forces of the conditions (such as terrain, coastlines, currents, weather, and infrastructure) in parts of the world where they may have to fight. To maximize yield from the military perspective, forward presence should be concentrated in areas of the world that are both troubled by the risk of war and important to the United States. In other words, the prospective returns on investments in operational awareness and interoperability are greatest when and where there is a reasonable likelihood that these capabilities will actually be used in war. There is, obviously, little practical value to be gained by making substantial investments in interoperability or operational awareness in regions where the United States has no intention to fight because it has no important interests to defend.

The United States currently maintains high levels of forward presence in several parts of the world—Western Europe, the Mediterranean basin, Northeast Asia, and Southwest Asia. Numerous factors undoubtedly affect this allocation of forward presence resources. Alliance commitments are, of course, a main factor, but they are to a great extent discretionary. Treaties between allies do not ordinarily specify the exact levels of routine, noncrisis military forward presence that the allies will maintain in each other’s territory. Rather, the treaties typically commit their signatories “only” to commit forces when war erupts. The North Atlantic Treaty of 1949, the founding document for the world’s premier and longest-lasting military alliance, makes no mention of military forward presence. The treaty commits its members to defend each other against attack and thus authorizes—but does not require any specific level of—U.S. military forward presence during peacetime.
ECONOMIC INTERDEPENDENCE AND
THE REGIONAL APPROACH
The aphorism that “trade follows the flag,” if it was ever true, no longer holds, at least in the sense that U.S. businessmen do not now require diplomats or the military to pave the way before opening shop on distant shores. There are no exotic ports left for the fleet to open, no remote, fog-shrouded kingdoms for soldier-diplomats to intimidate into accepting Western merchants and merchandise. The businessman’s way has already been paved, at least as much as military forward presence is capable of doing so. It is not physical access that keeps entrepreneurs out of some markets but rather doubts about the opportunities in those markets relative to others. Today, and even more tomorrow, the flag will follow trade. In that light, military forward presence can be a way of nurturing the political-military stability that is essential to continuing economic growth and political-economic reform. It can also be a tool for maintaining access to existing markets. The best places to apply that tool are the regions with which the United States is, or is becoming, economically interdependent.

Economic interdependence can be measured through statistics on trade flows and foreign investment. Thus the critical regions would be those that account for the highest proportions of American trade and foreign investment. All other things being equal, a region representing 20 percent of U.S. trade and investment ought to be accorded a higher priority in terms of military forward presence than one accounting for only 2 percent. Of course, “all other things” never stay equal over time; levels of presence should be adjusted to reflect the risk of instability in the high-priority regions. When the risk of instability in a high-priority region is low, relatively little military forward presence need be maintained. When the risk increases, consideration should be given to deploying forces.

Political-military alliances are, of course, important—but alliances are means to an end, not ends in themselves. It is well to remember this distinction between ends and means. No alliance lasts forever, and even when the form of an alliance remains intact, its substance may change. This has already happened to Nato. Once a mutual defense organization integral to the security of the United States, Nato has become a regional policeman without a central role in the defense of its North American members. Some of the positions taken by the United States in the mid-1990s appear to concede this point, at least indirectly. At the time, Clinton administration spokespersons correctly argued that even though the Cold War was over, membership in Nato was still in the national interest because it gave the United States “a seat at the table” of Europe. That is, it gave Washington an official platform from which to participate in European deliberations. In the mid-1990s the crucial European deliberations did not involve mutual defense
issues; they involved the humanitarian crisis in Bosnia and the economic matters associated with Europe’s nascent single economic community. Since the Clinton White House was (as the preceding administration was, and as the succeeding administration is) reluctant to participate directly in peacekeeping operations in Bosnia, it seems reasonable to conclude that the “seat at the table” was considered important primarily because it provided a way to ensure that U.S. economic interests were protected as Europe formulated community-wide economic policies and regulations.

With what geographic “units,” then, should U.S. policy makers deal, if even Nato no longer defines a zone sufficiently cohesive and homogeneous in terms of American interests? In general, it is useful to think in terms of regions smaller than a continent, or than the lands encompassed by an alliance like Nato, or than the expanses of territory that have been assigned by the Unified Command Plan to each of the military’s regional commanders in chief (or “CinCs,” such as of the Southern or Central Commands). All such groupings are too broad and heterogeneous; for example, the Southern Cone of South America differs markedly from the rest of the continent in terms of its economic growth, the robustness of its democratic reforms, and the absence of active insurgency movements and border disputes. On the other hand, focusing on units as small as the individual nation-state is too cumbersome, and in any event, some economic relationships and diplomatic-security obligations are transnational.

For these reasons, a regional approach is most suitable for the purposes of formulating requirements for forward presence. The regional “unit” further commends itself in that it would force strategists to look at priorities in new and different ways and thus avoid a pitfall all too common in long-range planning—implicitly assuming that the future will closely resemble today. Exactly how the regions should be defined will, of course, be a subject for debate; wherever the boundaries are drawn, incongruities and artificialities will result. The main point here is that geographical regions larger than states and smaller than such massive zones as continents or “CinCdoms” are useful units of analysis for determining U.S. national interests.

For the purposes of illustrating the regional approach to U.S. national economic interests and to the measurement of forward presence requirements, this article will focus on the Mediterranean Basin and South America.

The Mediterranean has been a major focus of American military forward presence, particularly naval, for nearly fifty years. Although the level of presence

The aphorism that “trade follows the flag,” if it was ever true, no longer holds, at least in the sense that U.S. businessmen do not now require diplomats or the military to pave the way.
has been reduced since the height of the Cold War, the U.S. Sixth Fleet currently maintains eighteen to twenty ships in the Mediterranean and conducts several dozen military exercises annually with local navies. Since individual ships rotate in and out of the Sixth Fleet on six-month cycles, between thirty and forty ships actually gain operational awareness of the Mediterranean and improve their interoperability with regional navies each year. For its part, South America is useful for the purposes of comparison, as that continent has traditionally received considerably less attention than the Mediterranean. For instance, the UNITAS exercise series, the principal naval forward presence activity in South America (annual exercises and port visits over a four-month period), involves fewer than half as many U.S. ships, aircraft, and personnel as are engaged in twelve months in the Mediterranean.

TRADE AND INVESTMENT

The economic interdependence between the United States and various parts of the world can be at least approximately measured through statistics on trade flows and foreign investment. The International Monetary Fund compiles data on the total value of U.S. exports to and imports from each nation in the world. These data on bilateral trade flows can be aggregated for geographic regions or for any other grouping or category of states. The same can be done for statistics on the amount of U.S. investment in every other nation and the amounts that other nations have invested in the United States.

By these standards, two geographic regions stand out as being of major economic importance to the United States: Western Europe (Ireland, Great Britain, France, the Low Countries, and Germany) and Northeast Asia (Japan, China, and South Korea). For some regions the absence of economic importance to the United States is equally obvious. Central Africa (Rwanda, Burundi, Tanzania, Uganda, Congo, and Kenya) is an example; the United States has no significant investment in or trade with that region. Most regions—including South America and the Mediterranean—fall between Western Europe and Central Africa in terms of the level of economic interdependence with the United States.

In the early 1970s, the entire Mediterranean Basin—that is, the countries with Mediterranean coastlines—accounted for approximately 10 percent of all U.S. trade (exports and imports combined). Since then, the Mediterranean’s percentage has steadily declined. During the last five years for which data are available (1994–98), its share has been in the 6.8–7.3 percent range. The high end of the range represents a 25 percent reduction over nearly three decades, the low end a drop of fully one-third. These reductions in the relative position of the Mediterranean region reflect the dramatic increases that have taken place in U.S. trade with other regions, particularly the Pacific Rim and the other parts of
North America. These data suggest that despite significant increases in absolute values, relative to other parts of the world the Mediterranean region has become significantly less important in economic terms to the United States.

What is true for the Mediterranean is true as well for each of its subregions. The subregions are North Africa (Morocco, Algeria, Tunisia, Libya, Egypt), the Middle East (Cyprus, Turkey, Syria, Lebanon, Israel, and Jordan—which has no Mediterranean coast but is usually considered a Middle Eastern state), and southern Europe (Greece, Italy, Malta, France, Spain, and Portugal). Albania and the successor states of the former Socialist Federal Republic of Yugoslavia (Slovenia, Croatia, Bosnia, and Serbia-Montenegro) could also be considered as part of southern Europe. Since none of the latter states has substantial economic ties to the United States, their inclusion would have no substantial effect on the data for that subregion.

From 1991 through 1998, the Middle East and North Africa subregions accounted for only about 1 percent each of U.S. trade. This is not a post–Cold War phenomenon; these subregions have accounted for roughly the same proportion of American trade since 1970, apart from a blip in the late 1970s and early 1980s. Of the three Mediterranean subregions, only southern Europe represents a considerable amount of U.S. trade, about 6 percent. Yet that percentage was lower in the late 1990s than at any other time in the last thirty years. Over the course of the 1990s, southern Europe’s share of all U.S. trade decreased by about 15 percent.

Foreign-investment statistics tell a similar story. Most U.S. investment in the Mediterranean takes place in southern Europe. As a percentage of all American overseas investment, that in the North Africa and the Middle East subregions is negligible; each accounts for less than 1 percent of all U.S. foreign direct investment. With respect to southern Europe, the preponderance of U.S. investment is in three countries—Spain, Italy, and France. In 1990 about 9.5 percent of the total value of all U.S. foreign direct investment was in these three. Since then, the proportion has been steadily declining, to 6.8 percent in 1998.

Not only do the trade and investment statistics suggest that the Mediterranean Basin has become relatively less important to the United States, but the figures actually overstate the region’s economic standing. This is because the statistics assume that all trade with and investment in France, Spain, and Portugal can be considered “Mediterranean.” France and Spain are, obviously, both Atlantic and Mediterranean countries, and a considerable volume of U.S. trade
with and investment in those countries is “Atlantic” in character rather than Mediterranean. Portugal has no Mediterranean coastline, but it is a member of Nato and, perhaps for lack of a better alternative, has been considered Mediterranean by U.S. strategists; it is, for example, in the Sixth Fleet’s area of responsibility. The American trade with France, Spain, and Portugal that flows through Atlantic seaports should be set aside when estimating the significance of U.S. economic interests in the Mediterranean; U.S. investment in those nations’ businesses and infrastructure projects that are oriented toward the Atlantic or Western Europe should similarly be excluded. (An example would be a factory that U.S. investors build in northern France near the Channel tunnel so as to maximize its access to markets in England and the Low Countries.) The implication is that the United States may be able to accomplish its forward-presence objectives with France, Spain and Portugal through Atlantic-oriented, instead of Mediterranean, activities.

Figure 1 depicts the proportion of U.S. trade represented by the two sub-regions of South America. The data presented on this chart indicate that both subregions are more economically important to the United States in terms of trade than are the Middle Eastern and North African subregions of the Mediterranean. While not high relative to Western Europe and Northeast Asia, U.S. trade with each South American subregion is about twice as great as for either North Africa or the Middle East. Unlike those for the latter, the percentages for the Southern Cone subregions have increased during the 1990s. Until the very late 1990s, trade with the Andes-Caribbean subregion also tended to increase; the drop-off in 1998 may have been anomalous, a partial result of the civil war in Colombia (which will be discussed below). The Andes-Caribbean subregion of South America consists of Bolivia, Peru, Ecuador, Colombia, Venezuela, Panama, Guyana, Suriname, and French Guiana. The Southern Cone includes Chile and the MERCOSUR (common market) countries of Brazil, Argentina, Paraguay, and Uruguay. As with Middle East and North Africa, U.S. investment in the Andes-Caribbean subregion is relatively insignificant as a percentage of all investment. On the other hand, American investment in the
Southern Cone has in recent years been increasing, presumably in response to political and economic reform. The value of U.S. investment in the three largest Southern Cone countries (Argentina, Brazil, Chile) approaches the value of that in Spain, France, and Italy. As figure 2 suggests, the value of U.S. investment in the Southern Cone will soon, if recent trends continue, exceed that for southern Europe. Increasing investment can be a precursor of increased trade volumes. Thus one might expect U.S. trade with the Southern Cone to surge in the future. Indeed, the long-term prospects for both investment and trade appear greater for the Southern Cone than for any Mediterranean subregion. This may be in part a consequence of the fact that unlike southern Europe, the Southern Cone may only now be entering the “spurt” phase of the economic development cycle—when annual growth rates are typically at their highest and greater than those of mature economies. Furthermore, the Southern Cone has more abundant and diverse natural resources than southern Europe; it also has a larger population and higher population growth rates. By 2020, the combined population of Brazil, Chile, and Argentina is expected to grow from 220 million today to 260 million. The combined population of Spain, France, and Italy is predicted to remain virtually flat over the next twenty years, at 160 million. Italy is expected to experience negative population growth rate during this period. On the basis of these economic statistics, the priority for the purposes of forward presence of each of these five subregions would be as given in the table.

Assuming that the threat of war or instability were equal for all of these subregions, the most rational strategy would be to calibrate presence according to priority. Absent a crisis, there would thus ordinarily be roughly equal levels of military forward presence for southern Europe and the Southern Cone; very low levels would be maintained in the Middle East, North Africa, and the Andes-Caribbean. Because economic trends change gradually, policy makers can be reasonably confident that a region that is determined to be of

**FIGURE 2**

**U.S. FOREIGN DIRECT INVESTMENT (FDI)**

![Graph showing U.S. foreign direct investment in Southern Europe and the Southern Cone from 1992 to 1998.](image)

<table>
<thead>
<tr>
<th>Subregion</th>
<th>Priority</th>
</tr>
</thead>
<tbody>
<tr>
<td>Southern Cone</td>
<td>medium (but increasing)</td>
</tr>
<tr>
<td>Southern Europe</td>
<td>medium (but declining)</td>
</tr>
<tr>
<td>Middle East</td>
<td>low</td>
</tr>
<tr>
<td>North Africa</td>
<td>low</td>
</tr>
<tr>
<td>Andes-Caribbean</td>
<td>low</td>
</tr>
</tbody>
</table>

Published by U.S. Naval War College Digital Commons, 2001
high priority in 2001 will almost certainly still be in 2004 and 2005. Equally, the status of southern Europe and the Southern Cone as medium-priority regions would not be expected to change dramatically over the short term. Thus it makes sense to base multイヤer plans for routine, noncrisis forward presence on enduring interests. Adjustments can be made if threat levels increase.

SECURITY THREATS
Predicting where and when an international incident will occur is a highly complex affair. Threats can develop quickly. That is why the United States has long maintained a large and sophisticated intelligence apparatus to identify trouble spots around the globe and to evaluate continuously the prospects for war and instability. Definitive assessments are, of course, beyond the reach of this article, but it does seem reasonable to assume that the probability of war is currently quite low in Western Europe, a high-priority region for the United States. Instability appears more likely in another high-priority region, Northeast Asia. Although the risk appears to be diminishing as North Korea opens itself to the West, until such time as the risk is significantly diminished, a high level of forward presence in Northeast Asia may make a strategic difference in terms of preventing war.

Of the two subregions in South America, the one that is more likely to experience instability is the Andes-Caribbean. Colombia is already consumed with a violent, anarchic civil war involving at least two major insurgent groups. The disorder has had ripple effects in neighboring Panama, Ecuador, Venezuela, and Brazil, effects that could well worsen. Cross-border conflict between well-armed Colombian paramilitaries and the military forces of neighboring countries is a distinct possibility, as is collaboration between the Colombian insurgents and criminal or rebel groups in nearby countries. This could have dramatic long-term consequences for American policy in the Andes-Caribbean region, as political and economic reforms are still quite fragile there. Conversely, the risk of instability in the Southern Cone is quite low. Democratic and economic reforms appear to have taken root, the territorial disputes between traditional rivals Chile and Argentina have been resolved, and the subregional nations are increasingly working together to address common challenges. For example, in 1996 Argentina, Chile, Brazil, and Uruguay collectively pressured Paraguayan military officers into abandoning a planned coup against the elected government.10

In the Mediterranean, tensions between Israel and its neighbors continue to flare, but there are no indications that the tensions will lead to another general war in the subregion. It is, in fact, difficult to determine exactly what benefits military forward presence can bring in the Middle East that are not already being achieved by U.S. diplomats in sponsoring the peace process—more recently,
attempts to restore the peace process—between Israel and the Palestinians. Furthermore, instability in the Middle East may simply be of the type that military forward presence is least likely to deter; that at least is what the historical record suggests. The Lebanese civil war of the early 1980s was only temporarily interrupted, not deterred or stopped, by the presence of the U.S. Navy offshore, or even by the “boots on the ground” of American, French, and Italian peacekeepers. Similarly, the intifada and the continuing violence that began in late 2000 in Gaza and the West Bank were obviously not deterred by U.S. naval forward presence in the eastern Mediterranean. Indeed, it may simply be unrealistic to expect Palestinian factions or extremists of any kind to forgive past grievances, relinquish claims for territory, or back away from their convictions simply because foreign warships routinely visit local ports, or because one indigenous navy or another can demonstrate its interoperability with the U.S. Navy in offshore exercises.

It might be argued that due to the proximity of the Middle East to Southwest Asia, military forward presence in the Middle East contributes to deterrence in the Persian Gulf. If this is true, the opposite must also hold—that is, that the objective of deterring interstate conflict in the Middle East can be served by military forward presence in Southwest Asia.

As for North Africa and southern Europe, at present the risk of war also seems to be low. In southern Europe, most of the major interstate issues arising out of the dissolution of “greater Yugoslavia” have now been resolved, about as well as anything short of total war can resolve them. Furthermore, so long as military peacekeepers from Nato, with or without U.S. participation, remain on the ground in Kosovo and Bosnia, the contribution that other forms of military forward presence can make to deterrence seems marginal.

As noted previously, however, there is more to forward presence than deterrence or demonstrations of national interest. Forward presence also enables the U.S. military to acquire operational awareness—practical knowledge about conditions in a theater. It also allows U.S. forces to improve their ability to work together with indigenous military forces and with allied forces that are not resident in the particular subregion but operate there.

Plainly, the benefits of current operational awareness of in-theater conditions and confidence in interoperability with friendly forces are time-sensitive; they are obviously most valuable when there is a high likelihood of war and when the United States is very likely to become directly involved. Just as plainly, operational awareness and interoperability are less valuable when conditions are

Indeed, the challenge for national strategists is to determine whether the regions that are important to the nation today will be equally important tomorrow.
pacific. They are least valuable when the United States has no significant interests to defend.

In subregions where high levels of military forward presence are continuously maintained, the United States is in effect making a considerable investment in current operational awareness and interoperability with friendly forces. Setting aside national interests for a moment, that is prudent when the risk of war is continuously high (as it was during the Cold War) but profligate when the threat is thought to be low. It may even be excessive when low threat levels are assumed to be transitory, because most forms of military forward presence can be increased if and when threat conditions become more adverse.

Based on this brief tour of the security horizon in South America and the Mediterranean, it would appear that absent a new crisis the only zone where the threat warrants a higher level of presence than economic interests alone would dictate is in the Andes-Caribbean subregion.

STRATEGIC COMMODITIES
Overall trade and investment statistics may obscure the significance of strategically important commodities to the United States. This could cause national strategists to underestimate not only the impact of a supply interruption on U.S. economic interests but also the contribution that military forward presence can make to preventing interruptions. Oil is the standard example, one that has particular salience given the recent spikes in oil prices as well as the tensions in the Middle East subregion.

There are two general aspects to the oil equation, production and distribution. It seems reasonable to presume that forward presence by the United States would tend to deter invasions of oil-producing states. That is to say, the presence of U.S. warships, air forces, and ground troops in a subregion can contribute to interstate stability by creating at least the perception that the United States is primed to defend oil-producing states from attack. This argues strongly for high levels of military forward presence in and around the Persian Gulf, due to the huge volume of production that might be lost in a war, and to the high costs of evicting an invader—as in Operation DESERT STORM. This same argument could be used to justify forward presence in other oil-producing or potentially oil-producing areas; none of the Mediterranean subregions fall into this category. Developments in the Middle East affect oil production by influencing the political decisions on oil output by Arab states in other subregions, but none of the states in the Middle East subregion as we have defined it (Cyprus, Turkey, Syria, Lebanon, Jordan, Israel) is a significant oil producer.

An area that does fall into the category of a potentially important oil producer is the Andes-Caribbean subregion of South America. Venezuela is a major oil
producer and shares a troubled and occasionally contested border with Colombia. Colombia, in turn, has both oil reserves and a domestic insurgency problem that could destabilize the entire region. Ecuador also has oil reserves and has already experienced spillover from Colombia’s turmoil. (For example, in September 2000 an estimated fifteen thousand Colombians fled into Ecuador to avoid the fighting between the Colombian government and an insurgent force.)

Oil distribution and pricing should, perhaps, be approached differently than oil production, at least in terms of evaluating the contribution that military forward presence can realistically make. Extensive naval forward presence in the Mediterranean did not keep oil prices low or supplies high during the 1970s, when two oil embargoes led to economic recession in the United States and Europe. Nor have high levels of forward presence of naval, air, and ground forces in Southwest Asia and the Mediterranean kept prices from sharply rising in 2000. Indeed, it is unrealistic to expect military forward presence of any kind to prevent sovereign states like Saudi Arabia and Venezuela from deciding to limit oil production in order to raise prices. Iraq may be the best example of the limited effect that forward presence can have in this respect. There has been an extraordinarily heavy military forward presence in and around Iraq for the ten years since the end of the Gulf War; large parts of Iraqi airspace are regularly patrolled by U.S. aircraft—a particularly intrusive form of presence made necessary by the international community’s desire to protect minority groups inside Iraq. Yet even under these conditions, Iraq has manipulated its oil production in an attempt to inflate the prices paid by the West.

Oil distribution to world markets can, of course, be disrupted in other ways. A state could mine or blockade a critical sea-lane. For example, during the Iran-Iraq War in the mid-1980s, sections of the Persian Gulf were mined by Iran, and Iranian Revolutionary Guards used captured oil platforms to attack tanker traffic near the Straits of Hormuz. (The U.S. response was to provide military escorts for the tankers.) Continuous military forward presence might deter such actions, but that is an expensive approach to what is ultimately a constabulary function. In comparison to the difficulties and expense of reversing production problems caused by the conquest of an oil-producing state, reversing distribution problems created by mining and blockades should be easy. It should be within the capacity of expeditionary forces from the United States or of local military forces that, in the Mediterranean, are part of the Nato structure and are more robust than their equivalents in many other parts of the world.
The bottom line appears to be that maintaining a continuously high level of military forward presence is not an efficient or effective approach to the threat of oil distribution problems. There is a more effective and perhaps more efficient approach—the strategic petroleum reserve. The United States built a strategic petroleum reserve for the express purpose of cushioning the effects of interruptions in distribution and of sudden price hikes. Investing in larger petroleum reserves is a better hedge against distribution problems and price hikes than is military forward presence.

IMPLICATIONS

Until the global balance of military power changes—and perhaps even after it does change—the United States should measure its military forward presence requirements on the basis of economic criteria. Such criteria should be applied in a regional framework, so as to guard against nostalgic assumptions that the parts of the world that are important to the United States today will be equally vital in the future. The basic allocations would then be adjusted as conflict in particular parts of the world became more likely. The proposed approach implies that military forward presence resources may not presently be allocated in a way that properly reflects the emerging future.

The various levels of routine, noncrisis military forward presence can be thought of as a continuum ranging from continuously high to none. The United States maintains continuously high forward presence in Western Europe, Northeast Asia, and the Mediterranean. Forces permanently stationed in these regions conduct cycles of interactions with local militaries (for instance, in exercises, information exchanges, planning, port visits, and other, largely ceremonial, activities). At the other end of the spectrum are regions like Central Africa, where the United States routinely maintains no military forward presence. That is to say, military forces are neither permanently stationed in nor periodically deployed to Central Africa to demonstrate national interest, deter interstate conflict, acquire operational awareness, or improve interoperability with local forces.

Most other regions fall between these two extremes. The two South American subregions are examples. In the Southern Cone subregion, the level of military forward presence might be classified as intermittent—largely confined to periodic port visits, exercises, and information exchanges. Presence in the Andes-Caribbean is determined by U.S. involvement in the war on drugs rather than an overall strategy for the subregion.

If national interests were determined on the basis of economic interdependence, it would appear the United States should consider reducing the level of military forward presence it maintains in the Mediterranean Basin. The savings could be transferred to other purposes, including force modernization.
savings could also be applied to military forward presence in other regions, such as South America. Economic statistics suggest, in fact, that the Southern Cone will soon become as important to the United States as southern Europe—a Mediterranean subregion where particularly high levels of military forward presence are maintained—and that the Andes-Caribbean subregion has a particularly severe risk of instability.

NOTES


5. These subregions are consistent with the groupings that the Sixth Fleet has employed in organizing its presence activities in the Mediterranean.


8. Ibid.


Developments in South Asia during 1999 reminded the world that the nuclear tests undertaken in 1998 by India and Pakistan had ominously increased the danger of nuclear conflict. Pakistani occupation of territory in the Kargil and Dras sectors on the Indian side of the Line of Control in the state of Jammu and Kashmir precipitated the fourth sustained engagement between the armed forces of India and Pakistan in the fifty years since independence and underscored the violent potential of incompatible claims in Kashmir by the two countries. The coup d’état in Pakistan in October 1999 raised questions about governance, stability, and democratic pluralism in a country that had spent about half of its political life since independence under military rule. Late in the year, the hijacking of Indian Airlines flight 814 by Pakistani militants introduced into the conflict the specter of state-supported terrorism. Together these events created the climate for the 28 percent increase in India’s military budget that was announced early in 2000.

SIMULATIONS OF THE UNTHINKABLE

Early in this chain of events, shortly after the nuclear tests in May 1998, the U.S. Naval War College, in Newport, Rhode Island, undertook a series of simulations and “decision events” designed to examine the consequences of these developments. The project started from the premise that the tests had increased the possibility of weaponization, deployment, and use of nuclear weapons in South Asia. In the tradition of games that the Navy had conducted in Newport for more than a hundred years, the college’s Decision Strategies Department organized a series of exercises and discussions to explore the implications of these events for the stability of the region.

Ambassador Taylor chaired the Asia-Pacific Studies Group of the U.S. Naval War College in 1999 and 2000. A retired Foreign Service Officer, he works with the Naval War College’s Decision Strategies Department, where he has directed studies on South Asian proliferation and Korean futures. While in the Foreign Service, he held a variety of assignments, including Deputy Assistant Secretary of State and chief of mission.

Naval War College Review, Summer 2001, Vol. LIV, No. 3
of events that gathered experts from the U.S. government, academia, foreign governments, business, private voluntary and nongovernmental organizations, and military commands to react to a prepared scenario set in the year 2003. Asking people from diverse backgrounds and organizations to interact with one another as they grappled with the issues posed by a hypothetical scenario was intended to produce insights that might have eluded an individual researcher or a group working within a single discipline. In addition to playing roles in simulations, participants were invited to develop their views in seminars employing a combination of anonymous commentary (using a networked array of computers) and ordinary discussions. The scenario depicted conflict in South Asia escalating from civic unrest and terrorism to an exchange of tactical nuclear weapons; the events ranged in length from four hours to six working days.

MAJOR ISSUES
In the interest of learning how the players representing the U.S. government, other governments, and other actors would respond to the hypothetical events, researchers acting as game controllers presented, in various simulation settings, successive segments of the event scenarios and then gave players free rein to react as they believed their “character” would in a real situation. In seminars, the questions posed were deliberately broad and open-ended, such as, “How do you think your organization or other organizations would respond to the events you have just heard described?” In this manner, the organizers tried to avoid constraining responses, as well as to encourage maximum interactions among participants.

In the exchanges that ensued, several questions emerged as salient in event after event. What could the United States or other actors in the international community do to discourage an escalation of hostilities between India and Pakistan? Could the United States or any other third country use military power to affect the outcome of a conflict in South Asia? Inasmuch as India and Pakistan account for only a minor share of international trade, would the impact on the world economy of a nuclear war on the Asian subcontinent be modest? If the United States and other countries wished to help ameliorate a disaster involving millions of casualties from a nuclear exchange, would they have the capacity to do so? Recognizing that the human and economic costs of a nuclear war in South Asia would likely be enormous, should the United States and other countries invest resources and effort now to reduce tensions between India and Pakistan, and also share with them the means of preventing an accidental disaster? Could such things be done while maintaining the arms-length posture the United States had assumed toward India and Pakistan in the name of nonproliferation? Finally and fundamentally, is a future scenario in which violence between India
and Pakistan that escalates to a substantial nuclear exchange plausible enough to justify worrying about the other questions?

THE GAME SCENARIO

In the tradition of military gaming, a difficult scenario—the overall “scripted” background and the situational framework within which role playing and discussion of important issues was to proceed during the successive events—was elaborated. Developed in consultation with experts on South Asia from the U.S. government and academia, the scenario was intended to stimulate planners to address tough challenges. The designers stressed to participants in all the events that the scenario was to be considered as only one possible future and that it was not intended in any sense as a prediction of the most likely evolution of events. The scenario was essentially the same in each event, but it became more elaborate as the series evolved—especially after an event known as the International Game, when the actions of foreign players moved the narrative farther than had been envisioned.

The scenario events began with a crisis postulated in the year 2003:*

The defeat of a new resolution in the United Nations Security Council calling for international involvement to resolve the status of Kashmir precipitated violent anti-Western and anti-Hindu demonstrations in India and Pakistan. This civic unrest was accompanied by a sharp increase in Islamic guerrilla activity in Kashmir. That activity culminated in the downing, by a shoulder-fired surface-to-air missile, of an Indian transport aircraft carrying the home minister, the minister of defense, and the army chief of staff, who were en route to Srinagar for an inspection visit. “Informed reports” implicated the government of Pakistan in the shoot-down. India responded by launching Operation RESOLUTE SWORD, air and artillery attacks against targets in Kashmir and northern Pakistan suspected of harboring and supporting perpetrators of violence in Kashmir and the rest of India. Simultaneously, the government of India declared that its intentions were limited in both scope and objective. Further, it issued an ultimatum demanding the immediate delivery of terrorist leaders who were sheltered in Pakistan, the dismantling of known terrorist headquarters and training facilities, and the removal of all Pakistani military forces from Kashmir. Initially, Pakistan offered little resistance to the Indian attacks, which inflicted heavy damage to the infrastructure targets against which they were delivered.

* Throughout, descriptions of hypothetical situations, actions, and policies posited by the game scenario are set in italics.
When Indian forces suspended their attack and began to celebrate the success of RESOLUTE SWORD, the Pakistani high command seized the opportunity for a surprise attack against Indian forces east and south of Lahore. During a two-day battle, Pakistani units managed to push about fifty kilometers into Indian territory, inflicting heavy casualties on Indian civilians, before a counteroffensive repulsed the Pakistani thrust. India matched its defeat of Pakistani troops in the Indian Punjab with a rapid movement across the Thar Desert toward the Indus River. Fearing that India was about to cut Pakistan in two, severing Islamabad’s economic lifeline to the south, the Pakistani high command ordered a barrage of tactical missile strikes. Four of these missiles carried nuclear warheads: three twenty-kiloton weapons were delivered against Indian forces to halt their advance toward the border, and the fourth was used against the supporting rail hub in Jodhpur. The strikes stalled the Indian movement and destroyed the rail hub but also caused widespread destruction among the civilian population of Jodhpur. Experts from the U.S. Defense Threat Reduction Agency estimated the number of dead and seriously injured in the hundreds of thousands.

INTERNATIONAL PERSPECTIVES

In January 1999, researchers organized a two-day international simulation to gauge reactions to these hypothetical events and the capacity of the international community to prevent, manage, and resolve such a conflict. The simulation brought together a multinational cross section of diplomats, academics, analysts, and military personnel. The countries represented were Australia, Canada, China, Finland, France, India, Iran, Japan, Pakistan, Peru, the Philippines, Russia, Singapore, the United Kingdom, and the United States.

Many of the diplomats present had reached the rank of ambassador in their countries’ diplomatic services; one had served as foreign minister. In the game, they were given roles as their countries’ principal representatives in the UN Security Council. In fact, they also had to replicate their governments’ entire decision-making authorities, because they worked without the benefit of instructions from home, and—with the exception of India and Pakistan—their actions in the simulation were “free play,” based on their individual best judgment of what their respective governments would do in response to the hypothetical situation as it unfolded. The scenario prescribed most of the military actions of India and Pakistan; however, even those nations’ representatives, who were experienced diplomats and scholars, devised their own diplomatic activities.

The International Game was unusual in having countries represented by their own nationals. This arrangement brought a greater degree of reality to the responses of foreign countries than is usual in U.S. government–sponsored simulations, which are generally played exclusively by American experts.
The participants were briefed on the scenario’s assumptions about the underlying situation for 2003. It assumed that unsatisfactory economic conditions had fomented significant unrest in both countries, leading to a rise in nationalist fervor and rhetoric. Both India and Pakistan had signed and ratified the Fissile Materials Cutoff Treaty and the Comprehensive Test Ban Treaty.* India’s real-world unilateral policy of “no first use” of nuclear weapons was incorporated in the scenario; Pakistan (reminiscent of the real-world posture of Nato during the Cold War) had made no such commitment. Both India and Pakistan had nuclear warheads that could be delivered by either aircraft or missile.

Move One of the simulation asked the players to react in the Security Council to early events in the scenario, including the civic unrest and stepped-up guerrilla activity, the shoot-down of the aircraft carrying Indian officials, and the beginning of Operation RESOLUTE SWORD. The Security Council moved swiftly into action. The Canadian delegate, in consultation with the belligerents, proposed a resolution calling upon India and Pakistan to cease hostilities and immediately disengage their troops on both the Line of Control and the international border. The resolution invited the UN Secretary-General to strengthen the military observer group already in place and deploy it on both sides of the border to witness the called-for disengagement. However, since the resolution would have committed troops without putting in place any new confidence-building measures, several representatives of the Permanent Five members** expressed reservations, and it was not adopted. Some of the delegates suggested that the presence of nuclear weapons made a traditional peacekeeping operation inappropriate.

At this point, the game-control cell advanced the players to Move Two, in which the Pakistani authorities ordered a nuclear attack against Indian troops advancing toward the international border. The Pakistani delegate justified his country’s attack as purely defensive and stressed that it had been directed at strictly military targets. The Indian delegation withdrew from the Security Council, declaring that the time for diplomacy had passed. The U.S. delegate and several others condemned Pakistan for its use of nuclear weapons, while China insisted that the international community bore a degree of responsibility for the Pakistani action, in that it had neglected to ensure that a military balance existed in the region.

* The game series ended before the U.S. Senate declined to ratify the CTBT. ** China, France, Russia, the United Kingdom, and the United States.

The International Game was unusual in having countries represented by their own nationals. This arrangement brought a greater degree of reality to the responses of foreign countries than is usual in U.S. government-sponsored simulations.
The Permanent Five considered putting sanctions in place against Pakistan but declined to intervene militarily to stop the crisis, fearing that such intervention would only raise the stakes, perhaps even lead to World War III. The Pakistani representative reacted to this decision with a mixture of disbelief and dismay, asserting that this conflict was World War III.

While the Security Council cell debated, the Indian player let it be known that his government had responded to the Pakistani strike by launching twelve nuclear weapons of its own, in an attempt to destroy Pakistan’s entire nuclear research, production, storage, and delivery capability. The game controllers assessed that it was unclear whether this objective had been fully realized, but inasmuch as many of Pakistan’s nuclear-related facilities postulated by the scenario were near populated areas, expert analysis by the U.S. Defense Threat Reduction Agency estimated some fifteen million casualties from the attack, including one to two million dead.

The Permanent Five, after separate consultations with India and Pakistan, now developed a proposal for ending the crisis: immediate cessation of military activities, renunciation and elimination of nuclear activities by both countries, a return to the status quo ante in Kashmir, and international security guarantees for both India and Pakistan. Several delegates opined that the security guarantees would be “difficult to sell” to their governments.

In a seminar discussion after the simulation, the senior Pakistani delegate explained that his government had been cognizant throughout the event of Pakistan’s unfavorable military balance with India. In an unscripted simulation, it would have sued for a cease-fire during Move One. The Pakistani players had been motivated in the game crisis, they explained, by eagerness to engage outside powers, especially China, Iran, and the United States, in forging a solution. The Indian delegate argued that his government would have worked very hard to reach a peaceful settlement but that once Pakistan fired nuclear weapons, only one response would have been acceptable to India. The Russian player, a scholar with strong ties to the Boris Yeltsin government (then still in power), explained that his country might see a nuclear crisis in South Asia as an opportunity to expand the Russian role in a system in which Russia was frustrated by its lack of clout. If breaking down the current nuclear proliferation regime could restore Russia’s proper voice in the international system, he believed, a Russian government might be willing for that purpose to renounce the Nuclear Non-Proliferation Treaty.

ECONOMIC PERSPECTIVES
In March 1999, a group of eighteen Americans—senior officials responsible for U.S. economic policy, business executives, academics, and military
officers—met for an evening and the following workday to examine the possible economic and commercial consequences of a conflict in South Asia. The participants either occupied or had recently left offices ranking from under secretary to deputy assistant secretary in the Departments of Commerce, State, and the Treasury. They were not asked to play roles per se but to offer their considered judgments about how the world would and should react to the events posed by the scenario.

As the scenario was introduced, and even before the downing of the Indian transport aircraft triggered a major escalation in the crisis, participants thought that an increase in tensions between countries with nuclear arsenals would cause investors and markets to jump to worst-case conclusions and precipitate an international financial crisis. Firms with economic interests in India and Pakistan would begin to hedge, checking on the security of their local employees and expatriates, evaluating financial exposure, and reviewing policy options. (The “real world” was to produce something of a reality check later, during the 1999 Kargil crisis, which would exhibit an eerie similarity to the scenario up to this point. Markets were to seem more relaxed in that actual event than the discussants had been, however, suggesting that participants in simulations may tend to anticipate escalation.)

When in the scenario military forces engaged in conventional exchanges, the discussants judged that international markets were likely to go into a tailspin, driving capital out of emerging markets to seek safe haven in the United States. Leading governments and international financial institutions would be pressed to resolve the resulting financial crisis. When the conflict widened, the participants judged, the humanitarian crisis within the subcontinent would increase the pressure on foreign governments for action. Business representatives strongly suggested that the most helpful way that governments could address the economic and commercial crisis would be to use every means available to defuse the underlying military and political crises before they deteriorated into a nuclear exchange.

When such an exchange occurred in the scenario, the participants concluded that a nuclear exchange of the magnitude postulated would create a humanitarian catastrophe. The situation would be hard for the government of the United States and other major countries to ignore, even though some people might say the governments of South Asia had caused the calamity and that it was up to them to deal with the damage. Relief efforts would also be complicated by

The scenario depicted conflict in South Asia escalating from civic unrest and terrorism to an exchange of tactical nuclear weapons; the events ranged in length from four hours to six working days.
residual contamination and the possibility of further conflict. Other questions would center on the capacity of the government of Pakistan to manage assistance and the willingness of the government of India to accept it.

Depending on weather conditions, the discussants predicted, severe shortages of food and potable water could exceed the capacity of relief organizations to respond and might even stress international markets. A conference of international donors would be required to mobilize the billions of dollars needed for relief. The prices of certain commodities, especially foods, could skyrocket and could trigger a global recession. Countless Indian and Pakistani citizens might seek refuge abroad; large numbers of refugees could cause other countries to create barriers to immigration. The most mobile people would carry away with them skills needed at home; the flight of elites might be matched by decisions of foreign investors in the region to take their chips off the table, resulting in profound and lasting damage to the economies of India and Pakistan.

OPERATIONAL PERSPECTIVES
More than fifty flag and field-grade American military officers, middle-grade and ambassadorial-level diplomats, civilian executives, and representatives of non-governmental relief and developmental organizations participated in a six-day simulation as part of Global 1999—the major annual simulation at the Naval War College. Military officers from Australia, Canada, and the United Kingdom joined them. Drawn from, inter alia, the U.S. Central Command, the U.S. Pacific Command, the Departments of State and Defense, the Agency for International Development, and the U.S. embassies in Islamabad and New Delhi, participants reacted to the unfolding scenario in ways that they believed their organizations would in a real crisis.

As the South Asian crisis began to develop in Global 1999, players pressed diplomatic measures to keep it from escalating. At the same time, nations effected voluntary withdrawals of their citizens from India and Pakistan. The eventual non-combatant evacuation operation was complicated physically by the distance of Islamabad and New Delhi from naval support ships at the outset, as well as by the dangers to military units inserted to conduct the evacuation. There was also a sense that the evacuation was sending a pessimistic signal as Washington was urging restraint on the local governments. The players did not anticipate that a nuclear exchange would occur as rapidly as the scenario dictated—only eight days after the downing of the Indian transport—so the evacuation was not completed before the nuclear strikes. Therefore, substantial numbers of Americans were presumed among the casualties, including many U.S. citizens of Indian origin, particularly those residing in the Punjab, from which they had been reluctant to depart.
While participants did not perceive a direct intervention or deterrent role for U.S. military forces in the crisis, they felt certain that in the aftermath of such a nuclear exchange any U.S. administration would look to the U.S. military, with its personnel, logistical, and technical resources, to carry heavy responsibilities in conducting decontamination, disaster relief, and early reconstruction efforts. Questions arose as to whether U.S. troops would be welcomed by the governments and the populations of the two countries, and whether the requirements of short-term crisis management might conflict with the longer-term interests of the United States in fostering constructive relations with both India and Pakistan. Military planners in the game sought to minimize the presence of U.S. forces on the ground and to define in advance the exit strategies by which responsibility would pass to civilian and nongovernmental leadership.

In the Global 1999 simulation, the physical requirements of disaster relief were found greatly to exceed current real-world preparations, which typically envision disasters on the magnitude of the bombing of the World Trade Center or the nuclear accident in Chernobyl. After the game, participants favored creation of a standing consequence-management force, an organization that could be deployed to alleviate disasters anywhere in the world. This force would be a ready and efficient alternative to current arrangements, under which regional commanders in chief are responsible for planning and organizing efforts, theater by theater. In its contingency planning, such a force would anticipate issues of coordination with other countries and determine how its activities would be directed on the ground.

Global 1999 players also concluded that although nongovernmental organizations were well equipped to assess local needs and deliver relief, they typically did not plan for major emergencies and certainly would not have the resources in personnel or materiel to contribute meaningfully in the aftermath of a nuclear exchange without enormous transfers from governments. Nongovernmental organizations would most likely have to perform “triage”—that is, explicitly identify the relatively small portion of victims they would be able to help, deliberately but unavoidably leaving many hundreds of thousands, perhaps even millions, without aid. Participants recommended more aggressive contingency planning and early integration of nongovernmental organizations into planning.

**IMPORTANT INSIGHTS**
This overall project examined a hypothetical problem from the perspectives of fifteen countries and specialties ranging from warrior to diplomat, missionary, and investor. Its most compelling result was that while participants differed on the details of their assessments and their remedies, and while some participants
in each event commented that the scenario pushed events more relentlessly toward nuclear war than they would expect in a real situation, no one argued that the scenario could not happen. In fact, they frequently expressed fears that it was a worrisome and consequential possibility.

The players in the simulations and seminars identified no significant military role for the United States or its allies in a military crisis between India and Pakistan. They asserted that though the United States and others might wish to help ameliorate a disaster involving millions of casualties from a nuclear exchange, the magnitude of resources and efforts required would be well beyond those envisioned in current contingency planning. The U.S. government, they suggested, should start now to create a standing, deployable “consequence management” force that could provide skills and material aid to help relieve the effects of major catastrophes anywhere in the world.

Participants argued that in economic terms alone, the costs to the world of dealing with the consequences of a nuclear exchange between India and Pakistan would far exceed the cost of trying to prevent one. On the other hand, they noted that the influence of the United States was limited and, in the simulations, only became more so as the crisis escalated. Still, time after time participants asked whether the international community was doing all it could to prevent terrorism over Kashmir and, indeed, whether more could be done to encourage India and Pakistan to reach a settlement in the Kashmir dispute itself.

An understanding of the problem needs to take account of the fact that India and Pakistan see different threats emanating from each other. Simply put, Pakistanis stress the injustice of India’s occupation of a large portion of a state in which Moslems are a majority, while fearing India’s stronger conventional forces; they accordingly seek international help to redress a wrong. Indians recall the accession of Kashmir by the governing maharaja at the time of partition—a legitimate procedure under international law, in their view—and reject any outside interference that could upset the status quo. Their main fear with respect to Pakistan is that its successful support of Kashmiri secession could cause groups elsewhere in India to seek new status for themselves—a prospect made worse by the specter of a fundamentalist Islamic government coming to power in Islamabad and mounting a jihad to free Kashmir. Any long-term solution will somehow have to give to Pakistanis a greater sense of security vis-à-vis their more powerful neighbor, and to Indians confidence that their pluralistic society can be protected against exploitation by outside agitators. The disagreement

The project started from the premise that the tests had increased the possibility of weaponization, deployment, and use of nuclear weapons in South Asia.
over Kashmir, which has provided the flash point for the series of conflicts since independence, might best be resolved as a secondary problem addressed in a context of broader concerns. That approach might also offer a way around the inherent conflict between attempting to apply outside pressure for resolution and simultaneously arguing that a solution has to be forged between the Indians and Pakistanis themselves.

Participants in the project’s simulations wondered if the United States and others could facilitate a solution in the same way U.S. diplomacy has contributed on several occasions to ameliorating conflicts in the Middle East and Northern Ireland, where conflicting claims have often appeared at least as intractable as those in South Asia. Some results had been achieved in those cases even though the governments of two close allies of the United States—Israel and the United Kingdom—had argued for years that for the United States even to talk to the groups that had terrorized their countries would be wrong and counterproductive.

Finally, many players throughout the simulation underscored the need to re-examine policies that weaken U.S. leverage in defusing potential crises. At present, as punishment for their proliferation, American policy denies military sales to India and Pakistan and the use of International Military Education and Training funds for attendance by their personnel in U.S. military schools and training courses. Additionally, military-to-military contacts with Pakistan are proscribed as a consequence of the coup d’état. American unhappiness with the nuclear tests and the coup needed to be expressed, but avoiding interactions with Indians and Pakistanis even on nonsensitive subjects has the effect of minimizing American influence on the very people who might push their countries into a military escalation, and it prevents the United States from getting to know the next generation of military leaders. In addition, current policy prevents the United States from engaging with the two governments on just the kind of confidence-building measures needed now on both sides to decrease the chance of a nuclear exchange: reliable nuclear command and control systems, for example, and nuclear threat reduction centers of the type the United States and the Soviet Union established, as well as cooperative threat reduction measures to reduce tensions between conventional forces along the border.

American policy, these simulations suggest, has stressed the laudable objective of discouraging nuclear proliferation to the point of dangerously reducing its own ability to discourage the use of the nuclear technology that India and Pakistan now possess. The United States would wield more influence over India and Pakistan if it accepted the fact of their nuclear status and attempted to establish safeguards and counter-use regimes, even while working to dissuade other countries from gaining a nuclear capability. In the last analysis, the Newport games warn, a policy based on the hope that either India or Pakistan is going to
abandon its nuclear arsenal is almost certainly wishful thinking and provides no basis for exerting U.S. influence in the urgent and difficult work of keeping those weapons from being employed.
“SECURITY IS LIKE OXYGEN”
A Regional Security Mechanism for West Africa

Lieutenant Commander Seth Appiah-Mensah, Ghana Navy

Since the end of the Cold War, the geostrategic significance of Africa to the United States has markedly declined, resulting in the treatment of Africa as a “backwater” in official U.S. policy making in recent years.¹ The derogation of African issues in U.S. foreign relations became evident as early as 1989, when war broke out in Liberia, a country hitherto regarded as having a long-standing special relationship with the United States.² But Africa’s, even Liberia’s, low priority in the dawning era failed to draw U.S. military intervention “to nip the civil war in the bud.”³ This prompted the Economic Community of West African States (ECOWAS) to form and insert an ad hoc military force—the ECOWAS Cease-Fire Monitoring Group, or ECOMOG—into Liberia in 1990.⁴ Initially designed for a brief operation in Liberia, ECOMOG has since deployed in two other states as well, Guinea-Bissau and Sierra Leone. Given the current chaotic, even hopeless, situation in Sierra Leone, and the less serious but still nebulous state of affairs in Liberia, Guinea, and Guinea-Bissau, subregional leaders have been under pressure to transform ECOMOG into a permanent regional force within a general ECOWAS security framework.⁵ To give effect to that dream, in October 1998 the ECOWAS Authority of Heads of States and Governments endorsed the establishment of a collective security regime known as the Mechanism for Conflict Prevention, Management and Resolution, Peacekeeping, and Security. (In this article, “West African subregion” refers, unless otherwise specified, to the ECOWAS community of nations.)
Though some outside observers see these regional initiatives as offering opportunities for African countries and their external partners, the new security arrangement poses many challenges. The West African subregion may not yet have the political, military, or economic means to accomplish the strategic objectives it has set for itself. In fact, this is a tall order; to imagine that ECOWAS can shoulder the burden alone would be a strategic miscalculation. If the security mechanism is to take off at all, let alone have any real chance of success, external support from the United States and other Western countries is crucial. Therefore, the United States and the international community may have to rethink their policies, reorienting them toward Africa in a way that reflects the current security dynamics in the region.

This article begins by defining the concept of security within the context of the volatile African environment. It then offers insight into the framework of the ECOWAS security architecture, as envisaged by its current leaders. Against the background of the political, socioeconomic, and military realities in the subregion, the article makes a case for strong external support to help bring about the desired dividends of the security mechanism.

Any critical analysis of the security paradigm in Africa requires a firm grasp of the unique African concept of security. For Africans, "emancipation is the freeing of the people (as individuals and groups) from the physical and human constraints which stop them [from] carrying out what they would freely choose to do. . . . Security and emancipation are two sides of the same coin. Emancipation, not power or order, produces true security. Emancipation, theoretically, is security."7

Thus, African security encompasses the whole range of dimensions of a state’s existence, including some that are already assured in developed countries. Generally, but especially in some states, such issues as internal disorder, the danger of food shortages, human-rights abuses, unequal opportunity, tribal imbalance in government and military institutions, insufficient social development, economic dislocation, colonial and neocolonial cleavages (mind-sets), and threats to freedom of speech constitute security problems in Africa.

**ECOWAS AND THE SECURITY FRAMEWORK**

Founded in 1975 as a forum for economic promotion and development, social and cultural matters, and the general progress of the continent, ECOWAS has emerged from the Liberian civil war as Africa’s foremost economic, political, and security grouping.8 ECOWAS adopted security-related protocols in 1978 and 1981, but none had been implemented at the time of the Liberian conflict.9
Thus, the conflict served as a wake-up call for the community to fashion its own security agenda. This conflict threatens the stability of the subregion—it has spilled over to neighboring Sierra Leone, and it poses serious threats of rebel attacks at the borders of Guinea.\(^\text{10}\) Guinea-Bissau’s recent political turbulence may also be linked to the war. It was, then, to contain and manage the current crisis, and to build a degree of security and confidence in the subregion, that ECOWAS endorsed a new security mechanism for West Africa.

The security system is to have as its highest decision-making body a Mediation and Security Council.\(^\text{11}\) The council may convene “as often as necessary” in the performance of its five primary functions: authorizing political as well as military interventions; determining mandates and terms of reference for such interventions; reviewing such determinations periodically; appointing such authorities as the special representative of the executive secretary and the force commander, upon the executive secretary’s recommendation; and informing the United Nations and Organization of African Unity of its decisions.\(^\text{12}\)

The council will operate through three committees. The Committee of Ambassadors will meet regularly and submit reports on regional peace and security issues to all council members, as well as affected states. Secondly, the Committee of Foreign Ministers, Defense, and Internal Affairs has the mandate to discuss the general political and security situation quarterly, or more frequently as necessary, and report to the council’s third committee, the Committee of Heads of State.\(^\text{13}\) In a provision characteristic of African organizations, an appointed Council of Elders would arbitrate, reconcile, and mediate during conflict.\(^\text{14}\)

The executive secretary will maintain a database of qualified and competent individuals who can serve on the Council of Elders. The executive secretary individually has been given an especially important role in conflict prevention and management—even to the point, innovatively, of deploying fact-finding and mediation missions on his or her own initiative. (The secretary will, however, be required to report any findings to the Mediation and Security Council.)\(^\text{15}\)

Another major innovation is a subregional Peace Observation System to provide ECOWAS early warning, and thereby opportunities for early action, to help the organization prevent situations from degenerating into violent crises.\(^\text{16}\) The idea emerged from the caucus of French-speaking African states that opposed and frustrated ECOMOG activities in the early 1990s;\(^\text{17}\) apparently there has been a change of heart and attitude among the francophone states—a cause for cautious optimism.

At the operational level, it is envisaged that a brigade-sized formation, to be called the ECOWAS Peace Monitoring Group (ECOMOG, as before), would
be set up as a permanent peacekeeping force. This “bank of rapid reaction force” (that is, a reserve of rapid-reaction capability) will comprise contingents from member states—earmarked, trained together, and organized for deployment at short notice. Of course, to train a force of this size will require a sizable and reliable logistical capability. As would be expected in an army-dominated organization, the organizers seem to have been oblivious to the invaluable role played by naval and air forces in Liberia and Sierra Leone, and to their future utility. This is a grave omission; naval forces from Nigeria and Ghana have been organic to ECOMOG since its inception and have proved themselves indispensable.

Patterned to some extent after the Nato security structure, this collective-security mechanism will require enormous amounts of administrative support (staff and offices), logistics, military resources, and funding, all of which would be difficult to obtain from Africa. Funding from the international community will therefore be crucial to this security project. However, to show its commitment, ECOWAS has proposed a community levy to fund its activities: each member state would be assessed 0.5 percent of the value of its imports from outside the subregion. The levy’s main rationale is easing the financial burden that contributing states would face during military operations.

U.S. POLICY TOWARD AFRICA

A prominent Western analyst has argued, “Africa is poorly understood by U.S. policymakers, [who] completely ignore the continent until some sort of politico-military crisis grabs their attention.” This encapsulates the nature of the attention that African issues have received from the United States, and from its major partners as well, in recent years. U.S. engagement in Africa lacks clarity and sense of purpose. Indeed, ever since the Berlin Wall came down it has been axiomatic for the United States, its Western allies, and the international community to pursue African policies that merely respond to crises rather than attempt to shape desirable outcomes.

Consequently, the amount of U.S. assistance to Africa has declined dramatically. From fiscal year 1985 to 1994, military assistance declined from $279.2 million to $3.8 million, and economic support funding from $452 million to fifteen million. For the same period, the total aid flow to Africa declined from roughly 10.3 percent ($1.87 billion) of an overall foreign-aid budget of $18.13 billion in fiscal 1985 to 7.6 percent ($1.36 billion) of $17.99 billion spent on foreign aid in fiscal 1994. This trend continued through the Clinton administration, which restructured the aid package according to four new general categories: sustainable development, humanitarian assistance, building democracy, and promoting peace.
However, the emphasis placed on democratic and economic development tends to exclude security factors, which are not only of vital importance to Africans but critical to their survival in the twenty-first century. As Joseph Nye has put the point, “Security is like oxygen—you tend not to notice it until you begin to lose it, but once that occurs there is nothing else that you think about.” Regrettably, such oxygen is ever in short supply in West Africa. To make even a start at reversing this trend, ECOWAS, in cooperation with external players, needs to optimize all the major components of state making. The three main components with respect to security—political, socioeconomic, and military—are so linked, interdependent, and intertwined that they can be regarded as a “trinity”: a weakening of one part drags them all down (although the triad is not a zero-sum game—increasing the importance of one factor does not decrement that of the others). It is no accident that more than three decades ago, Robert McNamara (who had just left office as the U.S. secretary of defense) expressed similar sentiments about security in developing nations: “As development progresses, [and] security progresses[, ]...[the people’s] resistance to disorder and violence will enormously increase.”

ENGAGING THE UNITED STATES
The essential threats to the viability of third-world states today have been described in a vivid fashion:

The most important...are the lack of internal cohesion, in terms of both great economic and social disparities and major ethnic and regional fissures; lack of unconditional legitimacy of state boundaries, state institutions, and governing elite; easy susceptibility to internal and interstate conflicts; distorted and dependent development, both economically and socially; marginalization, especially in relation to dominant international security and economic concerns; and easy permeability by external actors, be they more developed states, international institutions, or transnational corporations. This indeed is the predicament of West African states. Many factors contribute to this situation, which is characteristic of nearly all the countries in the ECOWAS group.

The Political Factor: Statecraft in Africa
One set of problems includes the incomplete stage of state making at which these nations find themselves; another involves the timing of their entry into an established international system. A third concerns their collective colonial heritage: “Rather than encourage integration, the colonial administrators fostered fragmented loyalties and ethnic particularism as part of their mechanism for ensuring effective control of these colonies.” Together, these three kinds of
problems have produced African states that are weak, vulnerable, and insecure. Many ECOWAS members have therefore become candidates for “failed state” status, which has already claimed Sierra Leone and Liberia.\textsuperscript{[31]}

The United States and other international donor countries have for the past two decades invested in what they perceive as viable West African democratic institutions, but the subregion still remains politically volatile.\textsuperscript{[32]} Nevertheless, in spite of this goodwill, a contradiction exists in U.S. policy making: “African countries, even if they do adopt political reforms, are unlikely to receive greater amounts of resources from a shrinking foreign aid budget.”\textsuperscript{[33]} Since recipient nations possess weak economies and no credible military might, the least armed resistance by a disgruntled politician or soldier causes their fragile political institutions to tumble.

**Socioeconomic Considerations**

Current statistics on the economic performance of sub-Saharan Africa reveal a dire situation, one that has the potential to degrade even further. The debt burden remains insurmountable and unsustainable. The region’s external debt in the 1990s averaged U.S. $204 billion, a per capita external debt of $359 for its estimated population of about 628 million (1998).\textsuperscript{[34]} External debt payments, which for Africa as a whole peaked at thirty-three billion dollars in 1997 (eight billion more than in 1995) represent more than 17 percent of the continent’s earnings from exports.\textsuperscript{[35]} The Secretary-General of the United Nations, Kofi Annan, observes that “this reflects a catastrophic fall in personal incomes on a scale seen by no other region since the Great Depression of the 1930s.”\textsuperscript{[36]} Perhaps it is time for Africa to have its own Marshall Plan.\textsuperscript{[37]}

Whilst this could be a long-term goal, the short and medium-term economic-aid objectives of the United States, other donor states, and international institutions need be synchronized and focused on specific critical areas. To that end, ECOWAS should be assisted to pursue and sustain collective prosperity through feasible economic reforms, macroeconomic coordination, and free-market practices.\textsuperscript{[38]} This would give ECOWAS a fair chance of survival as an economic entity, with a reasonable prospect of achieving the ambitious security architecture it has proposed for itself.

**Military Power**

The current military capabilities of the ECOWAS member states are inadequate to counter contemporary challenges and threats in the subregion. Even Nigeria, the subregional power, has struggled to maintain its presence in Liberia/Sierra Leone. As retired commodore Olutunde Oladimeji of the Nigerian Navy has
observed, “Nigeria has been operating with inadequate equipment [and] . . . budgets and [under] logistics constraints.” For this reason the new security mechanism envisions the pooling together of military strengths, in the form of an ECOMOG standing force that will be trained, equipped, and prepared to intervene in crisis areas. To this end, assistance in the form of equipment, funding, and training from the United States and other Western allies is critical. It would also be encouraging to see Nato initiate a dialogue with ECOWAS to share its invaluable experience in institutional arrangements and assist where necessary with equipment, military accoutrements, and technical cooperation.

Already, cooperative security arrangements exist between certain members of Nato and some of the ECOWAS states. The United States has initiated an African Crisis Response Force, training personnel in selected countries so that indigenous forces can deploy into crisis areas. Also, in 1996 ECOMOG enjoyed some $30.4 million in assistance from the United States. Britain and France are also pursuing similar but parallel initiatives in the region. It would be prudent to coordinate and synchronize all these efforts, if their full potentials are to be realized. This is very important, because, as one unofficial but authoritative study of U.S. government policy states, “unilateral U.S. peacekeeping or peace enforcement operations in Africa are not anticipated.”

**WHY SHOULD AMERICA BOTHER?**

It has been argued that for the United States, allowing African peace forces to resolve regional conflicts would be both a “load-decreasing” and a “dollarsaving” opportunity. Therefore, investing in an indigenous rapid reaction force in Africa will reduce significantly the demands upon American forces for intervention in Africa, and therefore, in due course, upon its operating budget.

Moreover, the current U.S. “low-priority,” risk-averse policy on Africa may ignore longer-term risks to its existing “investment.” During the early 1990s, for instance, the U.S. failure to intervene in Liberia (where more than 150,000 citizens were slaughtered) did not spare America the loss of facilities worth an estimated four hundred million dollars.

There is also a significant moral and cultural impetus for the United States to rethink its African policy. It is currently estimated that more than thirty million members (12 percent) of the population of the United States claim African-American heritage. These historical and cultural ties should be translated into viable political, economic, and military relationships in the new millennium. It will be interesting to watch whether the new American secretary of state, Colin Powell, allows Africa to be all but ignored in U.S. foreign policy.

Last but not least, Africa has over five hundred million people, as well as untapped natural resources. It is in the interest of the United States “to support
and preserve American values of freedom, individual rights, the rule of law, democratic institutions, and the principles of constitutional liberalism in the ECOWAS subregion. Regarding military aid, Washington has the "ability to influence African militaries to demobilize or downsize, reform and professionalize themselves, and engage more effectively in legitimate security and peacekeeping missions."

In his 1998 report on Africa, the UN Secretary-General restated, in effect, Robert McNamara’s 1968 observation (quoted above) on the relationships among development, security, and stability: "Peace and development," Kofi Annan declared, "remain inextricably linked—one feeding on the other, enabling the other and securing the other." Development has eluded Africans for far too long, as a result of nearly permanent economic crisis; it is not hard to fathom the reason why peace has also proven to be elusive.

To address the threats to security in West Africa, ECOWAS envisages a Nato-type security organization that would have an overarching responsibility for the security of the subregion. This is an ambitious but important security mechanism; however, inadequate military resources, fragile political institutions, and weak economies forebode great difficulty in getting started. Increased assistance from the United States, other nations and international organizations, and private institutions, properly coordinated and synchronized with respect to the desired "end state," should be pursued in the new millennium. In West Africa, the three factors of development, stability, and security are so inextricably linked as to defy piecemeal solutions. A holistic and comprehensive approach—one that engages the trinity of security development—is the key to unlocking the subregion’s security dilemma. In such a way, “the international community could help not only to support African governments and peoples, but also rekindle a sense of common purpose and shared humanity essential for planetary survival.”

NOTES
calculations of the United States regarding sub-Saharan Africa.


3. Ibid.

4. Initially composed of contingents from Nigeria, Ghana, Guinea, Sierra Leone, and Gambia.

5. Another reason for the establishment of the security mechanism was the tepid response of the United States and the United Nations to African crises.

6. See Eric G. Berman and Katie E. Sams, Peacekeeping in Africa: Capabilities and Culpabilities (Geneva: United Nations Institute for Disarmament Research, 2000), p. 379. These authors argue that African countries largely possess the troops and the will to intervene, but not the material means.


8. Berman and Sams, p. 79. The problems may be traced to several factors, including anglophone-francophone rivalry and suspicion that Nigeria harbors hegemonic ambitions.

9. Protocol on Non-Aggression, 1978; and Protocol Relating to Mutual Assistance on Defense, 1981; and Berman and Sams, p. 82. Neither the Defense Council nor Defense Commission was established, and no member state earmarked forces for the allied armed forces of the community.

10. Sierra Leone’s conflict has defied solution, rendering the country ungovernable. It is now a failed state.

11. The council will consist of nine members, to be elected for two-year terms. All decisions will require two-thirds majorities.


13. The Committee of Foreign Ministers, Defense, and Internal Affairs will comprise chiefs of staff, police chiefs, experts from ministries of foreign affairs, and representatives of immigration and customs services, border guards, and narcotics agencies. Its main function is to provide technical advice to the council. The Committee of Heads of State consists of the nine leaders of the member states of the council; it is to meet at least twice a year to deliberate on reports submitted to it.

14. The inclusion of the Council of Elders reflects a distinctive African value, whereby community elders are held in high esteem; they are the “reservoir of wisdom and knowledge” in dealing with difficult and sensitive matters.


16. This network would assess peace and security issues, including environmental, political, and social indicators. The subregion would be divided into three main zones: Ouagadougou (the capital of Burkina Faso), Monrovia (capital of Liberia), and Cotonou (in Benin).


18. Each mission will have its own strength, standard operating procedures, and rules of engagement.


21. A special fund is also proposed, based on voluntary contributions from member states and other friendly states.

22. Nigeria alone spent over ten billion dollars during the last decade on peacekeeping in Liberia and Sierra Leone.


24. The military-assistance figure does not include the roughly two billion dollars given to
Egypt annually. For economic support funds, see Schraeder, "Removing the Shackles," p. 194.

25. The euphoria that greeted democratization in Africa in the 1990s has waned, because rapid democratization of Africa has not stopped intrastate conflict. Neither has the uncoordinated economic aid that trickles into the continent.


28. Ayoob, p. 15.


30. Imobighe, p. 87.


32. Funding and observers from U.S.-based democratic institutions have facilitated multiparty elections, provided assistance to fledgling democracies, and promoted short-term visits by groups of Africans to the United States and by U.S. citizens to Africa—all to strengthen elements of civil society in both regions. Other donor countries, such as Britain and Canada, complement American efforts in this direction.


36. Ibid. In addition, under the William J. Clinton administration the United States adopted an approach to conflict and humanitarian disaster in Africa intended to contribute to regional stability. One of its aspects was financial support of operations carried out by other nations and international organizations. The second comprised assistance to African states in developing capabilities for conflict resolution; for subregional, regional, and international peacekeeping; and for humanitarian relief.

37. From 1947 to 1952, the Marshall Plan (named for Secretary of State George C. Marshall) helped to reconstruct European economies shattered by World War II. Similarly, postwar Japan received enormous American aid, the effects of which eventually also spilled over to other Asian economies. However, Africa—which was dehumanized through slavery and colonialism, and made to suffer proxy wars during the Cold War and post–Cold War conflicts—has yet to receive such attention.


41. INSS, 1997.

42. Ibid.


44. INSS, 1997, p. 164.

46. Strategy and Force Planning Faculty, p. 129.
47. INSS, 1995, p. 100.
One of the key debates that face naval aviation in the twenty-first century relates to its key equipment: the airplane. The United States is the only power that will be able to deploy a carrier force of any size, and it has held this position for over two decades. Soviet/Russian ambitions to deploy a blue-water air capability have been downsized, rationalised, abandoned, then reinstated under the threat of cancelation, and now lack funding. The People’s Republic of China is known to aspire to develop its naval aviation through the procurement of aircraft carriers, but it has made little obvious progress. While aircraft carriers enable the projection of airpower well beyond a nation’s shores without reliance on host-nation support, they have a major problem: they are expensive. Designing, running, and upgrading carriers are beyond the financial capacity of most nations. Only a few have the ability to deploy combat aircraft at sea, and the conventional aircraft carrier can only be procured in small numbers. While the United States can, within the politics of budget constraints, present a formidable air presence from carrier decks, the United Kingdom and France, the two middle-ranking powers with aspirations to maintain aircraft carriers, have been obliged by cost considerations to make some uncomfortable decisions.

Furthermore, there are continuing questions about the necessity of aircraft carriers for middle-ranking powers. It is argued that the aircraft carrier, by virtue of its considerable cost, is an unnecessary luxury. Under this scenario, the United Kingdom and France are perceived as being unlikely to embark upon independent naval operations but as
instead contributing to task forces dominated by the U.S. Navy, which in turn would provide the aviation assets. Practical experience, particularly for Britain, suggests that this view is dangerous.

Nonetheless, Britain’s experience of carrier aviation since the mid-1960s has not been altogether happy. A combination of a reduced world role and serious economic problems led to the downsizing of all British military services, with particularly savage cuts in the 1964–70 period. The aircraft carrier was deemed to be an expensive irrelevance. This view was shortsighted, ignoring the fact that Britain had a number of commitments and obligations in areas that had been brought into the ambit of Britain’s concerns through trade and colonialism. As always, plans failed to survive contact with the enemy—in this case, Argentina in 1982, when the only means of recovering the Falkland Islands (Islas Malvinas) was an amphibious operation. Since that time, despite ever-diminishing defence budgets, the aircraft carrier has returned to prominence, with the 1998 Strategic Defence Review (SDR) planning for two carriers capable of embarking around fifty aircraft, thus enhancing the deployability of British forces and increasing flexibility.

Nonetheless, there are a number of issues that need to be settled before the new vessels enter service in 2012. This article (based on the state of affairs in late 1999) seeks to provide a general outline of the options facing British naval aviation in the next ten to twenty years. It does not claim to be definitive but seeks to inform, highlighting in particular the manner in which the aircraft carrier has returned to the core of British military thinking as Britain adjusts to the conditions likely to pertain in the first two decades of the twenty-first century.

BRITAIN AND NAVAL AVIATION SINCE 1960: THE BACKGROUND
Decline has been a particularly notable factor in the United Kingdom, where cost considerations led to the abandonment of conventional-takeoff-and-landing (CTOL) aircraft carriers when HMS Ark Royal was retired in 1978. This was not a sudden decision.

The Decline of the British Carrier Force
The first threat to British naval aviation came in the 1960s. In 1957, a whole swathe of advanced aircraft projects had been canceled on the grounds of cost and a belief that their job would soon be done by missiles. Hawker-Siddeley Aviation’s P.1127, an innovative vertical/short-takeoff-and-landing (VSTOL) aircraft, was allowed to survive. The P.1127 was designed to prove the validity of VSTOL and was not intended for operational service. However, a derivative, the P.1154, was proposed by Hawker in 1962; it was finally abandoned in 1964, in favour of the F-4K Phantom.
It is arguable that the P.1154 had too much stacked against it. There appears to have been a suspicion amongst senior officers that a capable, supersonic VSTOL aircraft might have allowed a cost-conscious government to abandon a then-projected carrier design (known as the CVA-01) and replace it with smaller ships that could operate VSTOL air wings. The theory continues that the Admiralty had no intention of giving up its large carriers and agitated against the P.1154. This may or may not be true. What is not beyond dispute is that the Royal Air Force and the Royal Navy had completely different views on how the aircraft should be equipped; the naval P.1154 would have been a very different one from the RAF’s. In addition, the P.1154 was a technical risk. The P.1127 had not validated the concept of VSTOL at this point, and the notion of moving from a technology demonstrator with relatively simple equipment to a fully operational, supersonic type (the P.1154) was highly adventurous. Finally, there was already a superb, proven naval type in operational service—namely, the Phantom. The Phantom could be operated from the existing carrier fleet (with modifications), and it would obviously be fully compatible with CVA-01 from the start.

CVA-01, in spite of its innovative design, was vulnerable. The Royal Navy had five carriers (Ark Royal, Centaur, Eagle, Hermes, and Victorious) when Prime Minister Harold Wilson took office, plus two more (Albion and Bulwark) that had been converted into “commando carriers.” Of these seven vessels, only the Ark Royal and Eagle were large enough to accommodate the Phantom.

After a period of financial crisis, Wilson’s government decided to end the British military presence “East of Suez,” a conclusion that made the aircraft carrier an endangered species. In 1966, CVA-01 was canceled. While modifications to allow the operation of Phantoms from the existing vessels were financed, the carriers were not to be replaced. It was then decided that Eagle would not be modified to embark the Phantom. By the mid-1970s, it was clear that Britain would be left with only an antisubmarine helicopter force after Ark Royal’s withdrawal from service.

The argument put forward against complaints that this decision was shortsighted took the form that Britain’s reduced military commitments meant that naval operations where airpower would be required—opposed amphibious landings—would be coalition operations. The aviation assets required would come from the U.S. Navy’s carriers. Where the U.S. Navy was unavailable, the RAF would defend the fleet from land bases. This was a spurious idea, as the Royal Air Force did not have any suitable aircraft available for this task.
event of a war, the RAF would be busy elsewhere, and the Navy would have to hope that a U.S. carrier was nearby.

Consequently, the Royal Navy looked for alternatives. It became clear that the problem would be twofold. First, the government had to be convinced that a class of air-capable vessels was necessary; second, a suitable aircraft had to be found. The second problem was the easier to solve. The P.1127 had been developed into the Harrier, via a type known as the Kestrel. The Harrier entered service with the RAF in 1969. The U.S. Marine Corps placed an order (for the AV-8A) shortly afterward. Neither the RAF nor the Marine Corps envisaged using the Harrier as a fighter aircraft, even though Marine machines were wired for the AIM-9 Sidewinder.

The Royal Navy, faced with the choice of a navalised Harrier or nothing, decided upon VSTOL carriers. The major difficulty was to obtain the necessary vessels. Fortunately, in the late 1960s a new class of antisubmarine cruiser, the Invincible-class CVSG, was being designed. These already had a full-length flight deck and an island (for helicopter operations); Hawker-Siddeley Aviation was contracted to develop a minimum-change version of the Harrier to equip them. The modifications included the provision of radar and a raised canopy for better vision in air combat. The Sea Harrier entered service in 1979, some six years after the idea for its method of launching—off a “ski ramp”—had been arrived at.

The decision to develop a VSTOL fighter for the Royal Navy was not met with universal acclaim. Many felt that the Harrier offered little real military capability beyond being able to be based close to the front line without prepared runways. It was also noted that the proposed complement of aircraft for Invincible was less than half that of the Ark Royal. However, although the withdrawal of all but Ark Royal left the Navy with only around thirty fixed-wing aircraft, even this small force could have been useful. In the end, the only concession to demands to retain Ark Royal was the announcement that the name would be used for the third Invincible-class ship. This did not appear to be a sensible solution to the underlying problem of a capability gap. By 1980, the Navy had Invincible and also Hermes, which had been converted from the commando carrier role into an interim VSTOL vessel.

Four years after Ark Royal was retired, Argentina invaded the Falkland Islands. The only way to recover the Falklands was through an amphibious operation, and the Ark Royal would have proved invaluable. Although by comparison to a U.S. carrier air wing Ark Royal’s aircraft complement was small, a fixed-wing carrier would have provided the task force with strike aircraft, air defence fighters with a beyond-visual-range (BVR) missile capability, and, crucially, airborne early warning (AEW) aircraft. As the conflict progressed, it was clear
that all were desirable. As it was, the two aircraft carriers despatched (Hermes and Invincible) carried fewer combat aircraft than Ark Royal had done. No airborne early warning capability was available, and the most potent weapon carried by the Sea Harriers was the AIM-9 Sidewinder.

It is hardly necessary to recount that the Sea Harriers performed beyond all expectations, with not a single one being lost in air combat, while over twenty Argentine aircraft were claimed as shot down. In addition, the RAF version of the Harrier, then in its GR 3 incarnation, was successfully operated from the deck of HMS Hermes in the ground-attack role. The Sea Harrier and Harrier proved capable of operating in weather conditions that would have grounded other carrier-aircraft types, assisted by their ability to recover vertically onto the carrier deck rather than having to undergo the ordeal of arrested landings. Thus the concept of the VSTOL (more accurately, “short takeoff and vertical landing,” or STOVL) carrier was largely vindicated by the Falklands conflict. However, a number of problems were identified.

Lessons from the Falklands
The most obvious problem with the STOVL carrier was the lack of AEW aircraft. The lack of airborne early warning had made it possible for the Argentine pilots (air force and naval) to conduct daring attacks on the British task force, with minimal warning for defending combat air patrols (CAPs). As a result, the carriers had to be positioned farther from the combat area than was desirable. Further, the Sea Harrier’s light armament—two AIM-9s and two 30 mm ADEN cannon—gave it only limited combat persistence. This had not been a serious problem over the Falklands, but it was recognised that in different circumstances it might have been; after the war, this shortcoming was tackled with the provision of twin missile-launcher rails, doubling the number of Sidewinders carried. Finally, although the Sea Harrier’s Blue Fox radar performed beyond expectations, it was clear that a “look-down/shoot-down” BVR capability was desirable.

None of these problems would have arisen with the old Ark Royal: its Phantoms had had beyond-visual-range capability; the endurance of its combat air patrols had been extended through “buddy” refuelling (that is, from aircraft other than specialised tankers) by suitably equipped Blackburn Buccaneers; and airborne early warning had been provided by Fairey Gannet AEW 3s.

In the absence of conventional-takeoff-and-landing vessels, these lessons led to the development of the Sea Harrier F/A 2, which remains today the operational variant. The Sea Harrier F/A 2 is an impressive machine. Its performance in exercises has been remarkable. The F/A 2 is certified for the full range of attack missions, but the deployment of Harrier GR 7 (roughly equivalent to the AV-8B...
night-attack variant) from RAF units for attack missions is now commonplace, leaving the Sea Harriers for air-to-air operations. All RAF and Navy Harriers are now under a single command, Joint Force Harrier. With the Sea Harrier taking responsibility for air-to-air missions and the GR 7, with a larger weapon load, undertaking attack operations, the Royal Navy now has a suitable mix of aircraft for its carriers.

Experience has finally muted the criticism of the employment of STOVL vessels, at least in U.K. circles. The major problem facing the Royal Navy is that its carriers are too small. They cannot carry large air wings; they struggle to carry more than twenty aircraft. Thus, the Navy cannot have enough aircraft available on the scene of a crisis. The comparatively small size of its air wings reduces the number of sorties that can be generated, in turn reducing the carriers’ effectiveness. There are only two carriers in service at any one time (the third being in “deep maintenance,” or overhaul), which does not help matters. Consequently, although they are effective and have given good service, the Invincible-class carriers are not the best advertisement for STOVL types. The fact that the aircraft are good is obscured by the limited operations that can be carried out. This has led to the view that the STOVL aircraft carrier can in no way equal the versatility and flexibility of a CTOL vessel. To be allowed to exploit STOVL’s full potential, the Royal Navy required larger aircraft carriers. Defence spending policies in the 1980s and 1990s, however, meant that it was most unlikely to receive them.

Since the Invincible-class vessels are not due for replacement until around 2010, there has been until recently little consideration of what would follow them. With the end of the Cold War, British defence policy entered a period of confusion, as the government cut spending. Unfortunately, British military commitments did not diminish. Beginning with DESERT STORM, British forces became engaged in Iraq and Kuwait, as well as in the former Yugoslavia, protecting the safe haven established for the Iraqi Kurds, and patrolling the Iraqi no-fly zones. This was in addition to the tasks that they normally carried out. It became increasingly clear that asking British forces to do more with fewer personnel and less equipment was not a viable idea. This provoked the Labour opposition to promise that it would embark upon a Strategic Defence Review (SDR) if it were to be elected at the next general election. This duly occurred on 1 May 1997, and the enormous Labour victory meant that the SDR was likely to pass through Parliament without serious difficulties.

FUTURE AIRPOWER OPTIONS FOR THE ROYAL NAVY
The need for the United Kingdom to possess aircraft carriers has been questioned by a number of sources in recent years. A variety of critics have suggested that the British aircraft carrier is no longer necessary. Pundits in national newspapers
argued that the aircraft carrier is an expensive luxury and that the money needed for a new class of vessels could be better employed elsewhere.

The Strategic Defence Review and the Need for British Carriers

The SDR suggested otherwise. The review took longer to complete than had been anticipated in some quarters (as a result, it is rumoured, of objections from the Treasury), but it ultimately declared that the United Kingdom needed new aircraft carriers. The SDR changed the emphasis for British naval forces, moving away from the large-scale, open-ocean warfare for which it had been training for the past fifty years to force-projection and littoral operations in conjunction with the other two services. The SDR also made clear that the greatest importance would be attached to versatility and deployability.

It was obvious that the aircraft carrier would be integral to this vision. The SDR laid down proposals to procure “two large aircraft carriers capable of operating up to fifty fixed-wing aircraft and helicopters” from all three services. It should be noted here that what Britain considers a “large aircraft carrier” is different from what the United States does. For the Royal Navy, a “large” carrier has been the size of HMS Eagle or the old Ark Royal (53,000 tons fully loaded), able to carry between thirty and forty fixed-wing aircraft and helicopters. This is hardly “large” in comparison with the U.S. Navy’s aircraft carriers, since even the Midway (70,000 tons fully loaded) was larger than the British vessels. Still, the two carriers, scheduled to enter service in about 2012, will provide Britain with the ability independently to deploy combat forces to trouble spots. This is important, since there are reasons for Britain to deploy independently. Britain’s previous role as an imperial nation has left a legacy of close ties with former colonies. Some ten million British citizens live overseas, and there are thirteen dependent territories for which Britain is responsible. The need to be able to project military power to defend these interests or to provide aid to them is a clear reason for procuring newer, more capable carriers. In many instances, the infrastructure does not exist to support aviation operations in or around the overseas territories; the only means of sending aviation assets to such regions and then operating them is by means of a carrier deck.  

In addition to these direct responsibilities, Britain is a member of a number of international organisations, most notably the UN, Nato, and the Commonwealth. It is possible to envisage an attack on a Commonwealth country by a neighbour leading to British intervention, either on the side of the nation attacked or to impose or maintain a peace agreement. Although the SDR specifically denies that Britain seeks a role as world policeman, it asserts that there are compelling reasons for Britain not to be isolationist in outlook. The level of British overseas investment, particularly in the developing world (where British
investment amounts to the combined total of those of France, Germany, and Italy), demands that attention be paid to an ability to intervene in crises in these areas. Britain’s dependence on worldwide trade makes it essential for Britain to be able to defend its interests by the rapid projection of power to areas where conflict threatens to destabilise or damage them.27

SDR foresees the close integration of all three services: “By 2015, the Review expects further major change in modern warfare. Operations will no longer be characterised as land, sea or air. There will instead be a single battlespace in which land, maritime and air forces will be directed, targeted and supplemented by a new generation of intelligence, surveillance, information and communications systems offering a steep change in military capability.”28

Although this view presumes the development of new equipment and new technologies, the emphasis here is on the joint nature of future operations. This has already been manifested in the creation of the joint Harrier force.29 This is not all. The Royal Navy is adopting what is termed a “golf bag” approach to carrier-based aviation assets. The task force commander (the “golfer”) chooses the air assets (the “clubs”) needed to carry out the task at hand. The carriers deploy with the necessary mix of aircraft and helicopters. This might mean that the only helicopters carried would be the airborne-early-warning type, with antisubmarine helicopters being based on other vessels (for instance, HMS Ocean), if they were required at all. Thus the air wing might be made up of about sixteen Sea Harriers plus AEW helicopters. Alternatively, there might be a mix of Harrier GR 7s, Apache ground-attack helicopters, and a smaller number of Sea Harriers for operations where air defence was secondary to supporting troops. Additionally, the RAF’s support helicopter force (now amalgamated into a joint command of army, navy, and air force transport helicopter assets) has operated the Boeing Chinook from the Invincible class, so the carriers might be employed in a role akin to that of the American amphibious assault ship (LPH). The provision of a new carrier will make this “golf bag” approach even more viable, since it will make more room for the aircraft required for the mission.30

The issue now at hand is what type of carrier the new vessel should be. Neither the Strategic Defence Review nor any subsequent official paper has specified this in any way. Various proposals are being sought from industry, and it is unlikely that the type of vessel will be chosen until 2001 at the earliest.

The CV(F) and the FCBA

It is supposed that the “UK Future Aircraft Carrier,” or CV(F), will be of the short-takeoff-vertical-landing type, flying a “future carrierborne aircraft,” or FCBA, of some appropriate type. However, as far as is currently known, the Royal Navy could end up with a conventional-takeoff-and-landing vessel,
larger STOVL carriers, or what is known as a STOBAR ship—short takeoff but arrested recovery.

The CTOL type would appear to offer a number of advantages, especially in terms of the range of aircraft that can be embarked. Two suitable naval aircraft already exist—although not yet in frontline service—to equip this sort of carrier, namely, the Boeing F/A-18 E/F Super Hornet, and the Dassault Rafale M. The Royal Navy’s previous decision to depend upon “buddy” refuelling suggests that tankers would not be purchased and that strike/fighter types would likely constitute most of an air wing of forty to fifty aircraft.

Nonetheless, the CTOL carrier seems unlikely. Arrested landings are regarded as unnecessarily complicated by the Royal Navy after some twenty years’ experience of STOVL operations. STOBAR Options

The short-takeoff-but-arrested-recovery design has a number of operational problems. First, rapid flight operations are difficult unless an angled deck is employed. Even then, if the aircraft requires most of the flight deck to gain sufficient velocity to fly, aircraft will be unable to land while launches are taking place. A STOBAR carrier needs arresting gear, which, as noted, demands a large deck crew at a time when the Royal Navy will be seeking to embark as small a complement on the CV(F) as is viable.

Nonetheless, there is a reason for believing that the STOBAR option is not out of the running. The Eurofighter Typhoon is currently the subject of a British Aerospace study to assess its viability as a carrierborne aircraft, and it is quite clear that although the study is at an early stage, it is being taken seriously. In the past, attempts to convert land-based aircraft into carrier aircraft have not
been entirely successful, particularly with regard to stresses imposed on the airframe—especially in landings, inasmuch as carrier aircraft typically strike flight decks with greater force than land-based aircraft do runways. In the past, either the airframe has not proved strong enough or the weight increases caused by strengthening have imposed too-great penalties on performance.\textsuperscript{35}

In the case of a STOBAR Typhoon, the considerations are a little different. Structural strengthening would undoubtedly be required, but the weight increases may be minimised thanks to advanced technology. The Typhoon’s advanced flight-control system could be programmed to reduce the stresses of landing, particularly if integrated with a carrier-landing datalink. This would have a number of advantages. For instance, sudden pitching of the carrier deck would be recognised by the system, which would feed in last-second control corrections, ensuring that the aircraft landed within set limits. This would permit the airframe to be strengthened only as required for operations within those parameters—this, at least, is the theory that the British Aerospace study will investigate. There is little doubt that the use of thrust vectoring, already being planned for the Typhoon, coupled with a high-lift wing design, could provide near-optimal short-takeoff-and-landing capabilities for a “Sea Typhoon.” The use of a ski ramp would only enhance STOL performance.

There is another reason why a STOBAR vessel employing the Typhoon might not be out of the question—commonality. The Royal Air Force will be buying 232 Typhoons and has options for fifty more. The use of a navalised Typhoon would simplify servicing and lower unit-procurement costs. In addition, the United Kingdom would be the sole customer for the naval version; this would put design authority into the hands of British Aerospace. Experience with upgrading the RAF’s Jaguar strike aircraft (a collaborative project) has shown that with a single design authority that is not obliged to consult a partner, costs go down, and modifications arrive on time and on budget. In view of the chequered history of the Typhoon’s design and its increased costs, a cynic could contend that a navalised version unique to the United Kingdom, to be flown by a number of RAF units as well as the Navy, would prove extremely tempting to the Ministry of Defence.\textsuperscript{36} It would certainly enhance the capability of a joint force built around carriers. Reequipping the two Sea Harrier squadrons and the three Harrier GR 7 units (or an equivalent number of squadrons) with a navalised Typhoon would ensure that there were more than enough aircraft available for the CV(F).\textsuperscript{37}

---

One choice will not be available—to abandon the aircraft carrier entirely. Britain has tried this once, and for all the later success of the Invincible class, it had reasons to regret the decision.
Although it is currently unlikely that the Typhoon would ever become the sole combat type in the U.K. inventory, its use in the strike role as an early replacement for the Tornado is not an impossible scenario. This could make a navalised version more attractive, since it would slash servicing and training costs, thanks to commonality. There are a number of pitfalls with an air capability employing just one type; still, the point is that a STOBAR Typhoon might offer a number of cost benefits that a government concerned with defence spending would be hard pressed to ignore.

The Typhoon is not the only option. The Dassault Rafale M offers the advantage of being already available. There is little doubt that it is a capable airplane, and it has the swing-role flexibility offered by the Typhoon. Although the commonality aspect would be lost, the closer defence cooperation enjoyed between France and the United Kingdom since the mid-1990s suggests that the Rafale option would not face insurmountable obstacles. The RAF and French forces already run exchange programmes, and it is not impossible to envisage some form of joint operational-conversion unit, along the lines of the now-defunct Tri-National Tornado Training Establishment.

The precedents for such cooperation, however, are not altogether good. European defence projects have broken down over disputes about work sharing and requirements. The Typhoon programme, after all, has been severely delayed. The aircraft (as the Eurofighter) was meant to serve with the United Kingdom, Germany, France, Spain, and Italy; the vigorous disputes over which country should lead the design led to the project’s first acronym, FEFA (Future European Fighter Aircraft), being rendered in some aviation journals in unflattering ways. The concept-development stage of the Eurofighter saw the departure of the French, who built the Rafale instead. The political wrangling that ensued from increasing German opposition to the Typhoon delayed the in-service date for the aircraft dramatically; meant to achieve initial operational capability in the early nineties, the first Typhoon will not enter service with the Royal Air Force until 2002. The costs of the programme have risen drastically as a result.

While this experience suggests that a combined European defence force is unlikely in the near term, the idea of some degree of cooperation over CV(F) ought not to be ruled out. The Royal Navy could purchase a ski-ramp-equipped Charles de Gaulle-class carrier or two—although one hopes that any future vessels of that class will have power plants that function as advertised and decks large enough to accommodate their whole air groups. Although the Royal Navy seems not to want a nuclear-powered vessel, it would appear that the cost advantages (developmental and trials-related engineering expenses, etc., would be small) of adopting a STOBAR Charles de Gaulle-class ship might outweigh this objection. This would provide a European carrier capability; the Royal Navy
would foot far lower development costs, since these would have been absorbed by the French, and there would be commonality of aircraft types as well.

Consequently, the STOBAR option is plausible, with the choice between an existing aircraft type, the Rafale (which will be fully mature technology in 2012, but not outdated), and the slightly newer–design Typhoon, which is generally reckoned to be more capable. If the Typhoon proves suitable for carrier use, the decision between the two types would be difficult, but the difficulty would not be of the unwelcome sort. Whether, however, the short-takeoff–but–arrested-recovery option is the best for the Royal Navy is another issue. There is a compelling reason to suspect that Britain will procure another STOVL carrier—the Joint Strike Fighter.

**STOVL: The Preferred Option?**

Notwithstanding the scenarios in which a STOBAR vessel might prove attractive, there are strong reasons for the Royal Navy to continue to operate short-takeoff-and-vertical-landing vessels. This contention arises both from the nature of STOVL operations and from the United Kingdom’s involvement in the Joint Strike Fighter (JSF). The programmatic details of JSF are well known and beyond the scope of this article; the upshot is that although there are technological concerns, particularly with regard to the STOVL variant, cancelation seems highly unlikely. With its clearly projected costs for each version, the JSF is not an obvious candidate for the major overruns that lead to cancelation.

Furthermore, the Joint Strike Fighter is being relied upon to replace a huge number of aircraft, in a number of air arms. Outside the United States, it is the likely replacement for many F-16s and possibly F/A-18s. If the JSF is not procured, some other type will have to be. For the U.S. Marine Corps, the loss of the STOVL JSF—and in fact the programme is in jeopardy in the Defense Department’s current Quadrennial Defense Review—would be a serious blow, demanding either the updating of the AV-8B or a fundamental change in Marine aviation doctrine. This vulnerability prompted Lieutenant General Fred McCorkle, USMC, Deputy Chief of Staff for Aviation, to comment, “This nation has all its eggs in one basket.”

If the Royal Navy chooses a STOVL vessel, all of its aviation eggs will be in the same basket. It might therefore be said that the STOVL Joint Strike Fighter is a programme that will be made to work, since failure cannot be regarded as an option for two major customers. Even so, there are potential technical pitfalls with a supersonic STOVL type, and it is likely that the Marine Corps will be awaiting the flight-testing results with some anxiety. This will also be the case for Britain.

Britain has become a full collaborative partner in the JSF programme, and it has invested considerable sums in the project. It must be recalled that British
defence spending has edged inexorably downward for the last forty years; it is hard to envisage the JSF expenditure to date as a mere expression of interest in the project. This said, the cancelation of P.1154, TSR 2, F-111K, and Nimrod AEW 3 after prodigious expenditure makes procurement of the JSF less than a certainty.

Even given general confidence, however, that the Joint Strike Fighter will enter British service, a major concern remains: whether the STOVL JSF will work as planned. Supersonic VSTOL aircraft have been attempted before, and the competing design teams admit that the risk attached to the STOVL option is the greatest in the entire JSF programme. Apparently even simple and elegant technological and engineering solutions sometimes fail to work as advertised, and the inherent difficulties of short takeoffs and vertical landings make this more of a risk. Since the STOVL variant is only part of the programme, technical difficulties with it will not wreck the Joint Strike Fighter. Most customers will employ the CTOL version, and whether the STOVL variant works is not a concern to them. It is the U.S. Marine Corps and the Royal Navy that will suffer most if the STOVL JSF does not become a reality.

If the STOVL variant runs into development problems or escalating costs, the United Kingdom may well look for another carrier aircraft, employing the conventional Joint Strike Fighter in a few squadrons to meet its commitments to the programme. The problem here is that the choice of carrier platform will be made before the STOVL JSF is proven (unless it suffers major problems that instantly rule it out). If a STOVL platform is chosen with the intent to fly the Joint Strike Fighter, but the aircraft fails, the Royal Navy will have little option but to upgrade its Harriers. It may well be that attrition and the fatigue life remaining for the Harriers precludes an upgrade; there is also the issue of whether it would be better to use the Harrier GR 9 as the basis for an upgrade or to reopen the GR 9 production line. Commentator Roy Braybrook has argued that the U.S. Marine Corps and the RAF should not contemplate procurement of STOVL JSF but should instead rely on an upgraded Harrier GR 9/AV-8B with “improved radar and long range missiles.” In the event of STOVL JSF cancelation, it is likely that the GR 7/AV-8B would be used as the basis for the replacement type.

If the STOVL JSF does work, however, it will provide a number of clear benefits. It will offer the Royal Air Force the capability to operate from austere locations, even without runways. With an apparently increasing British commitment to peacekeeping—which usually occurs in areas where aviation infrastructure
either has been destroyed or never existed—this is important.48 For the naval element of a joint British force, JSF has several advantages over all other types in operations from the CV(F). The most obvious are that short takeoff and vertical landings remove the need for arresting gear, and for any launching system beyond the ski jump.

More generally, operations with the Invincible class have already demonstrated the manifest advantages that STOVL has over CTOL. These can be characterised as means of operability. First, because of the ski jump, a STOVL aircraft always leaves the deck on an upward trajectory, preventing a potentially dangerous lack of clearance between aircraft and ocean in rough seas with a pitching deck. Second, the likelihood of a “fouled deck” (from which aircraft can be neither launched nor recovered because others are in the way, either being moved or suddenly broken down) is far less acute. The vertical landing capability allows STOVL types to land even if their usual landing spots are blocked. There are no “bolters” (forced last-instant decisions of pilots not to land) on STOVL vessels; it has been proven that even when the carrier is blanketed in thick fog, with the pilot unable to see the vessel, landing is possible. During the Falklands conflict, one pilot returned from patrol to Invincible in just such weather conditions and nearly out of fuel. A searchlight was shone upward from the carrier, through the fog, and the pilot descended following the beam.49 Finally, the STOVL type can land on vessels that are not designed to operate it. This was taken to extremes in June 1983 by Sub-Lieutenant Ian Watson, who, through lack of fuel, could not return to the carrier; he instead landed on a passing Spanish merchant ship, the Alraigo. This saved an aircraft that would otherwise have had to be abandoned. It is not difficult to envisage situations where the ability to land aboard other vessels in a battle group would be advantageous.50 The United Kingdom has already commissioned HMS Ocean as an LPH, and that ship could operate a small number of STOVL types if need be.51

Although it has now become something of a cliché in British circles, the maxim “It’s far easier to stop, then land, than to land, then stop” has more than a grain of truth to it. It describes the manner of short-takeoff-and-vertical-landing operations perfectly. A 1992 report by the Center for Naval Analyses suggested that naval STOVL aircraft could undertake 25 percent more sorties than a conventional-takeoff-and-landing type over a five-hundred-nautical-mile radius in a twelve-hour period; if the radius of action were reduced to 250 nautical miles, the STOVL type could generate 40 percent more.52

To summarise, the STOVL vessel offers the ability to operate aircraft in weather conditions that would not be acceptable for CTOL types; it reduces the number of personnel required, by employing neither catapult nor arresting gear. Accordingly, the STOVL vessel can be cheaper and generate more sorties than a
CTOL carrier. For the Royal Navy, it is hard to see how there can be any objection to continuing to use STOVL vessels when these offer such great efficiencies compared to conventional types.

The major objection to STOVL appears to be that the Harrier and Sea Harrier are essentially limited in their technological advancement. Even this is debatable, however, since the technology of the Rolls-Royce Pegasus engine was advanced for its day; Rolls Royce has carried out continuous upgrading of and research on the engine, and further improvements are to be expected. As noted earlier, a key factor that is often overlooked when comparing the Sea Harrier and the more up-to-date AV-8B/Harrier GR 7 with other combat aircraft is that the former are fundamentally first-generation designs, employing 1950s engineering with 1980s modifications overlaid by 1990s technology. The only means of achieving adequate vertical and short-takeoff-and-landing performance without unacceptable weight penalties when the P.1127 prototype was designed was the single-engine vectored nozzle. This imposed certain limitations on the size of the aircraft, particularly on its internal volume. The Sea Harrier, Harrier GR 1/GR 3, and AV-8A/C, therefore, were all small combat aircraft. The AV-8B/Harrier GR 7 improved the airframe, avionics, and load-carrying capacity, but they were still tied to the vectored-nozzle system. Perhaps its most notable limitation, if the least relevant to operational efficiency, is that the vectored-nozzle engine does not allow supersonic flight. This has always received far too much consideration; supersonic performance has for too long been regarded as something that fighter and attack aircraft must have. The only two Western designs of note without it since the late 1960s have been the A-10 and the Harrier family. It is notable that these two types have always received only grudging respect (if any); even highly impressive results in combat have not saved them from verbal “friendly fire.” In the case of the Harrier, although General Norman Schwarzkopf announced that the AV-8B had been one of the key weapon systems in DESERT STORM, there were more than enough people prepared to criticise the AV-8B for its vulnerability to infrared-guided surface-to-air missiles rather than point out its effectiveness. While there is truth in this criticism of the AV-8B, it does not undermine the overall effectiveness of STOVL operations, and particularly not from aircraft carriers.

NOT AN OPTION BUT A NECESSITY

In light of all this, what is the best option for the Royal Navy for its future carrier operations? The conventional-takeoff-and landing carrier seems the least likely option, given the expense associated with the ship itself and the complement required to support the air wing. The choice between short-takeoff-but-arrested-recovery and short-takeoff-and-vertical-landing is more difficult to assess. If the
STOVL variant of the Joint Strike Fighter performs as planned, it will be extremely difficult to argue in favour of STOBAR. Still, the issue of how good an air-combat machine JSF is will need to be resolved before we can be absolutely certain that the Royal Navy will take delivery of the type. It should also be recalled that recent threats to the F-22 programme led to assertions that were F-22 procurement to be reduced (or cut entirely), the JSF would have to be redesigned to make it suitable for the air-superiority role. The Typhoon is claimed (by both its manufacturers and independent experts) to be the next-best air combat type after the F-22 (and less expensive), which may give it an advantage over rivals for procurement by the Royal Navy.

In spite of this, STOVL JSF would appear to offer a level of operational flexibility that neither a navalised Typhoon nor the Rafale could provide. The major difficulty with the Joint Strike Fighter is uncertainty whether the STOVL variant can be developed without major technical difficulties. We should know the answer in the not-too-distant future, although it appears that the programme may be subjected to some delay. If it works, the probability is that the JSF will operate from British carrier decks from 2012 as part of a joint Royal Air Force/Royal Navy force; STOVL is clearly the most flexible and affordable option for a power of Britain’s size. The difficulty is whether even the huge advances in technology since the P.1127 metamorphosed into the Harrier and Sea Harrier have been enough to advance STOVL from a proven but still developing system into a truly flexible weapon system. If not, defence officials will have to consider the options very carefully, and they will have to make some difficult choices.

One choice will not be available—to abandon the aircraft carrier entirely. Britain has tried this once, and for all the later success of the Invincible class, it had reasons to regret the decision, including a number of sunken vessels in the South Atlantic. This bitter experience makes clear the lack of carrier-based airpower would demand a fundamental shift in British foreign policy. There has already been such a shift since the end of the Cold War, and not toward less involvement. The current government purports to pursue an ethical foreign policy (despite some awkward contradictions); it has made this intention clear with its continuing support of action against Saddam Hussein, operations over Kosovo, and the deployment of troops to East Timor. Although in all three cases Britain’s contributions have not been the largest, they still represent a deployment of significant proportions of British military resources.

The Strategic Defence Review does nothing to suggest that there will be a scaling down of this support for humanitarian intervention and attacks against dictatorial states. Britain’s new foreign policy paradigm demands flexibility and
deployability. Even with the procurement of a strategic airlifter for the Royal Air Force (after much confusion), deployability will still be heavily dependent upon naval airpower. This is true now, and the CV(F) will be of immeasurable value when it eventually arrives. If the design of the vessel and the aircraft type to be operated from it are still far from clear, one thing is obvious. If Britain is to continue in the post–Cold War world in the role it has set out for itself, carrier-based air power is not an option. It is a necessity.

NOTES

1. See Derek Wood, Project Cancelled (London: MacDonald and Jane’s, 1975), pp. 207–24. The Admiralty did not appear to want the P.1154 and ensured that its requirements for the aircraft were almost diametrically opposed to the RAF’s needs.

2. Cynics could also suggest that the F-4K Phantom clearly required a CTOL carrier, thus ensuring that the new carrier would not be of the smaller VSTOL types. Less cynically, it could be pointed out that the Phantom was obviously a great aircraft already. To take the technical risk of procuring the P.1154 when there was a suitable aircraft available appeared pointless. Although the Royal Navy operated only one squadron of Phantoms at sea, this was a result of changing government policy and the abandonment of conventional carriers. Had the CVA-01 programme gone ahead with full replacement of then-extant carriers, at least five squadrons (and possibly six) would have been required.

3. These modifications included an extendable nosewheel strut to provide the correct angle of attack on takeoff, given the shorter carrier decks of the Royal Navy. This was mainly for reasons of the increased thrust provided by the Rolls-Royce Spey engine, but the political imperative of generating jobs in the British aviation industry demanded a high U.K.-built content for the F-4K. The Spey was also installed in the RAF version, in spite of the fact that the RAF did not want or need the engine; the Spey had a slower afterburner-lightup time and was less fuel efficient than the standard J-79. The net result of the Spey installation and the high amount of subcontracting led RAF pilots, however much they loved the aircraft, to describe them as the most expensive, noisiest, and slowest Phantoms in the world.


5. HMS Albion, Bulwark, Centaur, and Hermes were all of the Centaur class, with origins as light fleet carriers in the Second World War. There were substantial variations in each vessel, with Hermes often being regarded as a separate class by itself. In the early 1960s, Albion and Bulwark were converted to carry Royal Marine commandos, deployed by helicopter; they were roughly equivalent to the U.S. LPH in concept. Centaur was paid off in 1965 and used as an accommodation vessel. Of the fleet carriers, Victorious was the oldest, having entered service in 1941. It was modified extensively between 1950 and 1958, after which it was arguably one of the most advanced vessels in any navy. Again, in comparison with the carriers operated by the U.S. Navy, the Victorious air group was limited, with a maximum of thirty-six aircraft. Hermes could manage twenty-eight Sea Vixens, Scimitars, and Buccaneers; by the 1970s, however, only the Buccaneer was a truly viable combat type. Thanks to typically parsimonious defence spending, the Scimitar never received the avionics fit it needed to become a multirole aircraft. In its basic form it could carry four 30 mm cannon and early guided weapons, such as the Bullpup and AIM-9B.

6. Although Victorious might have managed to take Phantoms, more than thirty aircraft would have been a tight fit. Since at least ten aircraft would have been AEW Gannets and
Sea King and Wessex helicopters, this limitation would have reduced the effectiveness of the air group.

7. It was argued that this would be too costly, but it is difficult to find any authoritative source that accepts the validity of this reasoning.

8. The only aircraft types in the RAF’s inventory designed for such air defence work (apart from the Hawker Hunter, which had begun life as an interceptor but was by the late 1960s used for ground attack) were the F-4 and the English Electric Lightning. In the early to mid-1970s, the RAF’s Phantom fleet was committed to Germany in the strike role. After the introduction of the SEPECAT Jaguar, the Phantom was transferred to the air defence of the United Kingdom and thus would not be readily available to defend the fleet. In any case, how the RAF was to defend carriers in the middle of the Atlantic was never quite explained. The F-4 could not manage this, even with tanker support. Although the Lightning must rank as one of the great fighter designs (it gave an excellent account of itself in exercises against the F-14, F-15, and F-16 even in the late 1980s), its endurance was appalling. Designed as a point interceptor, even with in-flight refuelling the Lightning could not have maintained a combat air patrol over the fleet. In actuality, air defence of the Royal Navy was effectively left to the U.S. Navy—one of whose carriers, it was hoped, would be available to support operations.

9. This had been trialed by the United Kingdom, West Germany, and the U.S. Marine Corps in 1966, including landings on the USS Independence (CV 62) and small-deck tests on the dock transport USS Raleigh (LPD 1).

10. The CVSG had no arresting gear or catapults, and the Navy began to present it to politicians as a “through-deck cruiser,” which would operate an air complement of antisubmarine helicopters, with a few VSTOL aircraft (obviously Harriers) to defend against Soviet reconnaissance aircraft. The designation of the new class as a “cruiser” was barely plausible—the ship was cynically referred to as a “see-through cruiser”—and caused problems to begin with, since the type was initially meant to carry not fixed-wing aircraft but the Exocet antishipping missile to provide offensive firepower. Had the CVSG been designed to carry VSTOLs from the outset, it is likely that it would have been bigger, to give it a meaningful air group.

11. By 1972, although the decision to withdraw conventional carriers was not likely to have been reversed, it had been decided that a naval version of the Harrier would be acceptable. Such a machine, while not offering the capabilities of an F-4 or a Buccaneer, could at least provide limited air defence for the through-deck cruisers—now the Invincible class—while they conducted their primary mission.

12. There was considerable public support for the Ark Royal after a BBC TV series on the ship. Public feeling was so strong that the planned HMS Indomitable was renamed.

13. The Hermes was withdrawn from British service in 1985, after a period in reserve, and was sold to India, where it is now the Viraat.

14. The Ark Royal’s air wing normally consisted of twelve F-4 Phantoms, fourteen Blackburn Buccaneers, three Fairey Gannet airborne early warning aircraft, and six Sea King antisubmarine helicopters. Two Westland Wessex helicopters were also embarked in the search-and-rescue role, and occasionally a single carrier-onboard-delivery variant of the Gannet.

15. It might also be pointed out that the Royal Navy was subjected to savage cuts in the 1981 defence review. Under this, Invincible was to be sold to Australia. Had the Argentine leadership waited a few more months before invading, Britain would not have had the forces to mount the operation. Ironically, the man who came up with the ruinous review, Defence Secretary John Nott, was knighted in the aftermath of the Falklands campaign. It is arguable that without the review (which also called for withdrawing the patrol ship Endurance from the South Atlantic), Argentina might have judged that invading the Falklands was too much of a risk.

16. Later, after conversion to the commando carrier role, Hermes was fitted with a ski ramp, allowing it to operate Sea Harriers, thus providing an interim vessel for use while the three Invincible-class vessels were being built. The air wing on both carriers was reinforced, bringing the number of Harriers above the
twenty-six fighter and attack aircraft embarked by Ark Royal.

17. Thanks to the support of the United States, the Royal Navy was able to draw on Nato stocks of the AIM-9L all-aspect variant rather than use the then-standard AIM-9G.

18. The vertical takeoff requires far more fuel and reduces payload; thus it makes much more operational sense for the Harrier to begin with a short takeoff run.

19. The only solution was to modify a number of Sea King helicopters to carry the EMI Searchwater radar in a large, inflatable radome mounted on the side. The modification programme began while the conflict was in progress, but the first conversion did not arrive until the fighting was over.

20. The Argentine carrier, the Veinticinco de Mayo, was not employed. A planned strike on the British fleet from its deck was aborted for lack of wind; shortly afterward, the cruiser Belgrano was sunk by the fleet submarine HMS Conqueror. The Argentine fleet did not venture toward the Falklands again.

21. It was said that the British carriers were so far east of the Falklands that their crews qualified for the Burma Star.

22. It ought to be noted that the Argentine air force’s Mirage and Dagger aircraft had much the same problem. Even taking into account the fact that they had longer journeys to and from the combat zone, the air-to-air armament of the Mirage III and its variants was not particularly impressive.

23. Peter E. Davies and Anthony M. Thornborough, The Harrier Story (London: Arms and Armour Press, 1996), pp. 161–8. Although it is a potent machine, the Sea Harrier’s armament is limited. It invariably carries two 190-gallon fuel tanks under the inner wing pylons, as without them its range is prohibitively short. This leaves room for just two AIM-120s under the wings. Two more can be placed under the fuselage, but if they are, the 30 mm gun pods have to be off-loaded; for situations such as policing no-fly zones—where the ability to fire warning shots is useful—this is a problem. (This issue has been ignored for ridiculously small financial gains in the RAF’s forthcoming Typhoons, which the current government intends not to equip with a gun solely to save money on support costs.) The Sea Harriers usually fly with a mix of two AMRAAMS (advanced medium-range air-to-air missiles) and two gun pods in these circumstances, or replace the AIM-120s with the AIM-9M Sidewinder. The Sea Harrier can also carry ground-attack weapons, but the use of the Harrier GR 7 from carrier decks has seen a dramatic reduction in the use of the Sea Harrier for attack missions. Additionally, AIM-120s carried on the fuselage positions have been damaged by heat and vibration. At the end of the deployment of Sea Harriers in support of Operation ALLIED FORCE (the 1999 Kosovo campaign), around half the AIM-120s embarked upon Invincible had been rendered unserviceable by this cause.

24. The Harrier GR 7 is a very different machine from the Harrier of 1982. The “big wing” Harrier has a far greater load-carrying capability than the Sea Harrier, with eight underwing hardpoints (two dedicated for the use of Sidewinder), compared to the Sea Harrier’s four. As recently seen over Kosovo, the GR 7 can easily provide self-designation for laser-guided weapons, through the TIALD system, mounted on an under-fuselage pylon. The combination of Sea Harrier and Harrier GR 7 provides greatly increased flexibility for British maritime operations. The GR 7 (as well as the GR 5 model, which preceded it) suffered at first from a number of problems affecting the engine and weapons. One of these problem areas was the new 25 mm ADEN cannon, specially designed for the Harrier GR 5/GR 7; after years of trying to integrate it with the Harrier’s weapon system, it was decided to abandon the idea. The considerably less advanced 30 mm ADEN (based on the German 30 mm cannon designs of 1944–45) still functions perfectly, but it can only be carried by the Sea Harrier. All the surviving GR 5 aircraft have now been upgraded to the GR 7 standard.

25. This criticism usually revolves around the capabilities of the Sea Harrier. In spite of its upgrade, it must be recalled, the Sea Harrier is a first-generation STOVL type, based on the P.1127, which was never designed to carry a war load. The Harrier GR 7, as an evolved, second-generation type, gives better insight into the potentialities of STOVL for naval use.
26. It is not hard to posit such a scenario—for example, had Italy not permitted the use of its airfields for Nato operations against Serbia, the only nearby alternatives were either not Nato-compatible (e.g., in Hungary) or not adequate (as in Albania).


28. Ibid.

29. The Sea Harrier force is moving to the RAF base at Cottesmore to be in closer proximity to the RAF’s Harriers: two squadrons operate from Cottesmore and two more from the nearby base at Wittering, which is literally just a few miles down the A1.


31. The provision of catapults and arresting gear adds cost to the vessel, and the joint nature of future operations means training of RAF pilots to conduct carrier landings as an unlooked-for complication that would add expense, which many perceive as pointless when STOVL types have worked so well for the United Kingdom.


35. This factor (among others, of course) led to the adoption of the F/A-18 over the F-16 by the U.S. Navy; the F-16 was felt not to have the necessary development potential for carrierborne operations. The Royal Navy has experience with aircraft that are not strong enough for carrier operations. Although the Seafire (navalised Spitfire) was highly regarded, for instance, it was more prone to landing accidents and downtime than other types.

36. This is speculation on the part of the author, who is not entirely convinced that procuring navalised Typhoons would reduce the overall unit cost of the aircraft. The costs of developing the aircraft might well be greater than the savings made by having a single, national design authority.

37. The key issue here would be training pilots of the RAF squadrons to land on carriers, discussed above, but perhaps more relevant is the role portrayed for the Typhoon in RAF service. It has already been mooted as a contender—in modified form—for the Follow-on Offensive Aircraft System (FOAS), which will replace the Tornado GR 4. Defence spending priorities may well cause more Typhoons to be purchased to replace Tornados in some squadrons before the FOAS is chosen. Since the early 1990s, Ministry of Defence policy, even if unstated, has been to reduce the number of types in the RAF’s inventory. The Buccaneer and Phantom were withdrawn some years before schedule; the Jaguar fleet was constantly targeted for retirement until the Jaguar GR 3 upgrade added vastly improved capability at a very low price. It also demonstrated that requirements unique to one country enabled the problems of collaborative ventures to be swept aside with ease.

38. Currently, the RAF possesses the following types in the following number of combat aircraft squadrons (including operational conversion units): Tornado F3, six; Tornado GR 1/GR 4, eight; Jaguar GR 3, four; Harrier GR 7, four; and Sea Harrier, three. The Typhoon is scheduled to replace all the Tornado F 3 and Jaguar units. It is anticipated that the total of 232 will permit the formation of ten frontline squadrons and a training unit to teach both air-to-air and air-to-ground operations. If the Harriers were replaced with navalised Typhoons, the number of “operational conversion units” (OCU—equivalent to the U.S. Navy’s fleet readiness squadrons) could be reduced from the current five to two (one Typhoon, one Tornado).

39. A problem with this scenario might be whether one OCU for the Typhoon would be enough. The question of whether it would be wise to rely upon one aircraft type would also have to be considered.
40. This is to say nothing of the creation of a European defence force. A number of senior European politicians have called for the formation of European armed services. It should be noted that British politicians are not among them. It is probably fair to say that the British electorate’s respect for its armed services and its often suspicious attitude toward foreigners (at least those who do not speak English as their native language) make it unlikely to take kindly to the politician who initiates the transfer of British forces to supranational control. Aside from British insularity (or xenophobia), there is considerable anecdotal evidence that the idea might not work. It is not widely known, but some French units in Bosnia refused to speak anything other than French to their colleagues (this from a conversation between the author and a British officer who served in Bosnia). French spirits rose when a French Canadian unit arrived, but the differences between Canadian and “metropolitan” French were such that the Canadians decided that they would rather use English.

41. See the “Straight and Level” column in *Flight International* on many occasions in the mid-1980s for full explication of the term.

42. It is not unfair to say that the departure of Avions Marcel Dassault was met with relief from the other nations. Dassault has developed an unfortunate reputation for promoting its own products at the expense even of collaborative programmes with which it is involved. This led to the Aeronavale (French naval aviation) receiving the Super Etendard instead of the SEPECAT Jaguar M, which was arguably the better aircraft. See *World Air Power Journal*, vol. 35, p. 68, for particular reference to the Eurofighter Typhoon.

43. Ibid, pp. 54–97. The Eurofighter story became almost farcical. When it was announced that the aircraft was to be named the Typhoon, objections by the Germans (who felt it tattleless to name the aircraft after one that had killed many Germans in the Second World War) led to the compromise that it would be known as the Typhoon only for export customers. The RAF was thus faced with the prospect of having the “Eurofighter” as its main combat type.

44. The *Charles de Gaulle* suffered embarrassing power-plant problems (related to auxiliary engines, not the nuclear reactor) on its maiden voyage, being forced to return to port early. It was also discovered that the flight deck was not large enough for its E-2C early warning planes to manoeuvre safely. See “Carrier Concerns,” *Jane’s Navy International*, March 1999, p. 8; and “Up Close: *Charles de Gaulle* Finds Its Sea Legs,” *Jane’s Navy International*, November 1999, pp. 10–1.


46. The Harrier GR 7 is to be improved to the GR 9 standard with the addition of new avionics.


48. Operations in such cases, though, are unlikely to take the form of the 1970s and 1980s, when Harrier GR 3s operated from forest clearings using short runways of pierced steel planking, or from shopping-mall car parks. It is more likely that austerely equipped airports would be employed instead, with the STOVL capability removing the need for long, well-maintained runways.


50. See Davies and Thornborough, pp. 161–2.

51. Short takeoffs would not be possible, since one of *Ocean’s* close-in weapon systems would be in direct line of a takeoff run; VTOL would be the only option.


53. Supersonic performance might be given to the Harrier through the use of a technique known as “plenum chamber burning.” This has been under test for many years but has never been introduced for operational service.

54. There have, of course, been other, less well known, types, such as the Italian/Brazilian AMX (the obscurity of which seems a little unfair).

55. See “Raptor under Threat,” *Air Forces Monthly*, September 1999, p. 4. Also, see
“United States Navy Today: Part 1, the Carrier Air Wing,” *Air International*, August 1999, pp. 100–6, where Tony Holmes posits the view that “JSF is effectively a bomber with a strong fighter capability, more f/A than F/a.”


57. The SDR identified a need for a heavy-lift capability for the RAF and called for the interim leasing of “C-17 class” aircraft to fulfill the role. Much to commentators’ surprise, the strategic airlifter was canceled in 1999. Thereafter it became increasingly apparent, however, that UK defence commitments demand this type of aircraft, and the requirement was reinstated.
TRANSFORMING THE U.S. ARMED FORCES

Rhetoric or Reality?

Thomas G. Mahnken

The leadership of the Defense Department has enthusiastically endorsed the proposition that the growth and diffusion of stealth, precision, and information technology will drastically alter the character and conduct of future wars, yielding a revolution in military affairs. President George W. Bush campaigned on a pledge to transform the U.S. armed forces by “skipping a generation” of technology. A month after assuming office, he promised in a speech at the Norfolk Naval Base to “move beyond marginal improvements to harness new technologies that will support a new strategy.” He called for the development of ground forces that are lighter, more mobile, and more lethal, as well as manned and unmanned air forces capable of striking across the globe with precision.1

Secretary of Defense Donald Rumsfeld stated during his confirmation hearings that his central challenge would be to “bring the American military successfully into the 21st century.”2 Soon after assuming office, Rumsfeld commissioned Andrew W. Marshall, the Pentagon’s premier strategic thinker, to conduct a fundamental review of American strategy and force requirements. The review reportedly recommended that the Defense Department emphasize forces capable of fighting and winning wars in Asia, with its vast distances and sparse infrastructure, in the face of increasingly challenging threats.3

Speaking at the U.S. Naval Academy in May 2001, President Bush called for “a future force that is defined...
less by size and more by mobility and swiftness, one that is easier to deploy and sustain, one that relies more heavily on stealth, precision weaponry, and information technologies.” He also committed himself “to fostering a military culture where intelligent risk-taking and forward thinking are rewarded, not dreaded,” and to “ensuring that visionary leaders who take risks are recognized and promoted.”

The U.S. armed forces themselves have embraced—at least rhetorically—the need to transform so as to meet the demands of information-age warfare. They have fielded new capabilities, such as stealth and precision strike, and explored novel approaches to combat, such as network-centric warfare and effects-based operations. Nevertheless, significant organizational barriers to the adoption of new technology, doctrine, and organizations exist. The services have been particularly reluctant to take measures that are disruptive of service culture, such as shifting away from traditional platforms and toward new weapon systems, concepts, and organizations. The Army’s attempts to field a medium-weight ground force, the Navy’s development of network-centric warfare, and the Air Force’s experience with unmanned air vehicles illustrate such difficulties. In each case, efforts at transformation have faced opposition from service traditionalists who perceive threats in new ways of war. For the Defense Department to succeed in transforming the U.S. armed forces, it must both reallocate resources and nurture new constituencies.

THE CHARACTER OF WAR IN THE INFORMATION AGE
Recent years have witnessed the rapid growth and diffusion of information technology. It is radically changing the structure of advanced economies, the nature of politics, and the shape of society. It is also shifting the ways in which wars are fought. What many refer to as the emerging revolution in military affairs (RMA) is merely the military manifestation of the information revolution. The shape, scope, and strategic impact of the revolution is uncertain. Still, the experience of recent conflicts, together with trends in the development of technology, suggests changes in the conduct of war on land, at sea, and in the air, as well as the growing use of space and the information spectrum for military operations.

One trend that is already apparent is the ability to achieve new levels of military effectiveness by networking together disparate sensors, weapons, and command-and-control systems. Rapid advances in information and related technologies already allow military forces to detect, identify, and track a far greater number of targets over a larger area for a longer time than ever before. Increasingly powerful information-processing and communication systems offer the ability to distribute this data more quickly and effectively. The result is a dramatic improvement in the quantity and quality of information that modern
military organizations can collect, process, and disseminate. In the future, as in the past, forces that can secure a superior understanding of their own dispositions, those of their adversaries, and the features of the battlefield will be at a considerable advantage.\(^5\)

In a number of instances, the U.S. armed forces have attempted to explore how improvements in situational awareness can increase combat effectiveness. From September 1993 to September 1994, for example, the U.S. Air Force conducted an experiment that pitted eighteen F-15Cs equipped with Joint Tactical Information Distribution System (JTIDS) terminals against unmodified F-15s. JTIDS provided a datalink that allowed each modified F-15 to share its sensor and threat data with all the others. Their unmodified opponents were supported by E-3A Airborne Warning and Control System (AWACS) aircraft but could share information only by voice radio. The enhanced situational awareness provided by JTIDS allowed the modified F-15s to achieve an exchange ratio that was in their favor by a factor of around 2.6.\(^6\)

The increasing use of information technology portends a significant shift in the balance between offense and defense, fire and maneuver, and space and time. Militaries that harness the information revolution are already at a marked advantage in comparison to those that do not. The Gulf War hinted at the battlefield advantages that accrue to armed forces that capitalize on stealth, information, and precision weaponry. Nato’s air war over Serbia stands out as another demonstration of at least the tactical effectiveness of advanced military technology.

The integration of information technology into military forces is also changing the relationship between fire and maneuver. Networking long-range sensors and weapons allows us to concentrate fire from dispersed platforms on a common set of targets. The U.S. Navy, for example, has examined the “Ring of Fire,” a concept for focusing dispersed naval fire on shore-based targets.\(^7\) Networking thus allows the potential massing of effects without massing forces. It could also reduce vulnerability by denying an adversary the ability to target forces with his own long-range strike systems, while increasing the tempo of military operations by reducing the delay between observation and action.\(^8\) By operating faster than adversaries, a networked force may effectively deny them battlefield options.\(^9\) These trends favor networked forces that are small, agile, and stealthy over hierarchical organizations that are large, slow, and nonstealthy. Should the U.S. armed forces exploit these trends, the United States will gain increased...
tactical, operational, and—potentially—strategic leverage over potential adversaries.

While the United States currently enjoys a considerable lead in exploiting the information revolution, it is hardly alone in attempting to do so. Indeed, the list of militaries interested in information-age warfare is long and growing. Some may develop strategies to deny foes the ability to project power into their spheres of influence. Others may challenge the United States in space or the information spectrum. Moreover, their ability to do such things is growing. The director of the Defense Intelligence Agency, for example, has testified that Russia and China, as well as other smaller states and nonstate actors, are pursuing capabilities to disrupt, degrade, or defeat American space systems. Similarly, one recent article assessed that twenty-three nations have the ability to launch information-warfare attacks. Failure to meet such threats could lead to a military that is increasingly irrelevant to the types of wars that the United States will fight.

Past revolutions in warfare have changed not only the character and conduct of combat but also the shape of the organizations that wage war. The emergence of new ways of war has altered the importance of existing services, and combat arms triggered the rise of new elites and eclipsed previously dominant ones. During the first half of the twentieth century, for example, naval aviation assumed a central role in war at sea. As the aircraft carrier displaced the battleship as the centerpiece of modern navies, naval aviators challenged the traditional dominance of surface warfare officers. During the same period, the advent of land-based aircraft created new elites within armies and eventually spawned new military services. Armored forces usurped the roles of cavalry in armies across the globe. The information revolution portends similar organizational turbulence as the character of war on land, at sea, and in the air changes and as combat spreads to space and the information spectrum.

THE U.S. ARMED FORCES AND THE EMERGING RMA
The Department of Defense has declared its recognition of the need to change radically the structure of the U.S. armed forces in order to embrace the information revolution. The 1997 Quadrennial Defense Review committed the department to transforming its forces. As then–Secretary of Defense William Cohen put it:

The information revolution is creating a Revolution in Military Affairs that will fundamentally change the way U.S. forces fight. We must exploit these and other technologies to dominate in battle. Our template for seizing on these technologies and ensuring military dominance is Joint Vision 2010, the plan set forth by the Chairman of the Joint Chiefs of Staff for military operations of the future.
The congressionally mandated National Defense Panel argued even more strongly in favor of the need to transform U.S. forces. The panel’s report urged the Defense Department to “undertake a broad transformation of its military and national security structures, operational concepts and equipment, and . . . key business processes,” including procurement reform. It recommended, among other things, that the department accord the highest priority to a transformation strategy designed to prepare the United States to confront the new and very different threats of the twenty-first century. It also argued that the department should place greater emphasis on experimenting with a variety of systems, operational concepts, and force structures.

In 1998, the secretary of defense and the chairman of the Joint Chiefs of Staff designated U.S. Joint Forces Command (or JFCOM, formerly Atlantic Command) as the Defense Department’s executive agent for joint experimentation. Since assuming this responsibility, JFCOM has explored the concept of “rapid decisive operations,” including attacks against critical, mobile targets—a mission that places a premium on nearly simultaneous sensor-to-shooter data flows and high-speed, long-range weapons. The command plans to hold large-scale exercises to test new operational concepts in 2002 and 2004.

Beyond such initiatives, however, the Defense Department has yet to implement its announced commitment to transform its forces. The American armed forces today look much the same as they did ten years ago, only smaller. They have emphasized improving their ability to accomplish current tasks over exploring new ways of war. Similarly, most major acquisition programs of the last decade have represented incremental improvements to current systems. The services have fielded relatively few new weapon systems; of these, only a tiny fraction, such as the B-2 stealth bomber, could have major impacts on the conduct of war.

Advocates of transformation point to the need to shift from a force based upon major weapon systems to one based upon networks. They argue that precision-guided weapons, platforms to collect enormous amounts of information about the enemy, and command and control systems to direct one’s own forces will play increasingly important roles in warfare. While the services have invested increasing amounts of money in information technology, budget data on major acquisition programs suggest that the U.S. military services continue to have strongly platform-centric approaches to procurement. More than 75 percent of the Department of the Navy’s major-acquisition budget for fiscal year...
2002 is committed to large, traditional platforms—for instance, a new class of submarine (SSN 774), carrier-based aircraft (the F/A-18E/F), various surface ships (DDG 51 and LPD 17), and the tilt-rotor V-22 for the Marine Corps. U.S. Army and Air Force programs show comparable emphases upon platforms.\textsuperscript{18}

Rhetoric about transformation has yet to be reflected in weapons the services acquire, let alone the way they acquire weapons. The Army’s attempts to transform itself into a medium-weight force, the Navy’s experimentation with network-centric warfare, and the Air Force’s investment in unmanned combat vehicles all illustrate the difficulties associated with exploring new approaches to combat.

**The U.S. Army and the Medium-Weight Force**

The Army faces the challenge of transforming itself from a tank-heavy force designed to protect Western Europe from the armored columns of the Warsaw Pact to one capable of responding to contingencies worldwide on short notice. Operation \textsc{Allied Force}, Nato’s war against Serbia, highlighted the Army’s lack of units that are light enough to move quickly yet heavy enough to strike hard. The experience prodded the Army chief of staff, General Eric Shinseki, to launch an effort to reconfigure the Army into a more mobile yet still lethal force. In October 1999 he announced a goal of transforming the service into a “medium-weight” force capable of deploying a five-thousand-man brigade anywhere in the world within ninety-six hours. As he put it, “We must provide early-entry forces that can operate jointly, without access to fixed forward bases, but we still need the power to slug it out and win decisively.”\textsuperscript{19} He designated two brigades at Fort Lewis, Washington, as test beds for exploring new concepts and organizations. These units have traded in their tracked M1A1 Abrams tanks and M2 Bradley fighting vehicles for wheeled LAV III infantry fighting vehicles leased from Canada. They are also examining innovative new tactics and organizations. In November 2000, the Army awarded a four-billion-dollar contract to build the “Interim Armored Vehicle,” a new generation of light, wheeled vehicles with which to equip the new medium-weight units.

A key element of the Army’s transformation is the Future Combat System, a network of light—and possibly unmanned—vehicles that would replace tanks and self-propelled artillery in medium-weight units. Planners intend that the new vehicle will weigh no more than twenty tons (compared to the seventy-ton M1 Abrams), so that it can be transported aboard the Air Force’s most numerous transport aircraft, the C-130. Because it will lack the armor to slug it out with enemy tanks, its effectiveness will depend on its ability to identify and engage enemy forces before they can engage it.\textsuperscript{20} The Army’s plan for the Future Combat System is quite ambitious: the service plans to choose a design before
Shinseki leaves office in 2003; production is to begin in 2010; and the system is to be fielded by 2012. The General Accounting Office has, however, expressed concern that key technologies may not mature quickly enough to meet such a timetable.\textsuperscript{21}

The Army’s transformation plan is not without its detractors. The merits of a medium-weight force composed of wheeled vehicles remains to be demonstrated. Moreover, the prospect of a medium-weight force threatens the traditional emphasis upon armor as the centerpiece of ground combat, a notion that has defined the service for the past six decades. Indeed, it challenges the very definition and purpose of armored units. It is therefore hardly surprising that both active-duty and retired armor officers and enlisted men have been vocal in their opposition to the replacement of the tank with lighter wheeled vehicles. Many are particularly uncomfortable with the prospect of trading their heavily armored tanks for more vulnerable, if more mobile, vehicles.\textsuperscript{22}

Nor is it certain that the Army will maintain its current course. This is not the first time that the Army has attempted to transform itself. Indeed, it has examined the structure and organization of its combat units on twelve separate occasions over the last sixty years, accumulating a track record that is at best mixed.\textsuperscript{23} It remains to be seen whether the current effort will survive General Shinseki’s retirement.

\textit{The U.S. Navy and Network-centric Warfare}

The U.S. Navy faces the challenge of transforming itself from a fleet designed to fight in the open ocean to one that can dominate the littorals and project power ashore. Like the other services, it must also define its roles in space and cyberspace. To carry out these tasks, the Navy has sought to link weapon, sensor, and command and control systems—that is, to wage network-centric warfare. The Marine Corps, for its part, is exploring new methods of power projection and attempting to come to grips with the challenges associated with military operations in urban terrain.

The Navy’s track record of innovation is checkered. The demise of the Arsenal Ship highlights the barriers to innovation within the service. The Arsenal Ship, a vessel built to commercial standards and manned by a small crew, would have packed enough firepower to stop an armored column. Despite enjoying the support of Admiral William Owens (the vice chairman of the Joint Chiefs of Staff), Admiral Jeremy M. Boorda (the Chief of Naval Operations), and General Charles Krulak (the Commandant of the Marine Corps); the program lacked institutional support within the Navy. Critics raised questions about the utility and effectiveness of the ship. In addition, the ship lacked a constituency within the Navy. Indeed, it appeared to threaten a number of constituencies inside and
outside the Navy. Some surface warfare officers and aviators saw it as a threat to the aircraft carrier, while submariners saw it as stealing a mission they themselves wanted. Still others disliked the idea that the Arsenal Ship’s considerable firepower could be at the disposal of a ground commander. These communities attempted to undermine the case for the Arsenal Ship. Indeed, Admiral Boorda was forced to move the program from the Navy to the Defense Advanced Research Projects Agency in an attempt to preserve it. The ship’s opponents were aided by people in industry and Congress who had stakes in the status quo. As one former congressional aide put it, the Arsenal Ship “was a threat to the carrier, and that was a threat to Newport News Shipbuilding. And that, in turn, was a threat to the Virginia [congressional] delegation.”

In November 1997 the Navy killed the program, which a year earlier it had declared one of its highest priorities, due to “insufficient funds.”

At a deeper level, it appears that the Arsenal Ship challenged the Navy’s traditional notion of command. The vessel was essentially a truck designed to bring ordnance within firing range of targets. It would have lacked the sensors to target its own weapons, and it would have possessed only a minimal self-defense capability. Officers who had for years aspired to command destroyers, cruisers, and aircraft carriers likely did not relish the thought of becoming truck drivers.

Nonetheless, in recent years the Navy has begun exploring concepts that would replace large platforms with a network of smaller and less vulnerable systems. The Navy Warfare Development Command (in Newport, Rhode Island) and the Naval Postgraduate School (in Monterey, California), for example, have examined STREETFIGHTER—a family of small platforms designed to gain and sustain access to the littoral region in the face of a strong resistance, or “access denial”—as well as CORSAIR, a small aircraft carrier. Further, the Navy Warfare Development Command, stimulated by the performance of HMAS Jervis Bay in East Timor, is exploring the use of fast catamarans to deploy and sustain amphibious forces. Other Navy innovators have proposed converting Ohio-class SSBNs to carry special operations forces and large numbers of land-attack cruise missiles.

Such ideas have predictably drawn fire from officers who see them as a threat to existing surface ship programs. STREETFIGHTER in particular represents a challenge to the Navy’s current approach to force structure, which emphasizes a relatively small number of large, highly capable ships. Rather than conducting rigorous analysis of the benefits and limitations of such platforms, STREETFIGHTER’s detractors have tended to engage in ad hominem attacks. Vice Admiral Dan Murphy, the commander of the Sixth Fleet, was remarkably blunt...
in his criticism of STREETFIGHTER: It is “a wild idea. . . . There is nothing behind it. There is no analysis." You know, [Vice Admiral Cebrowski] dreamed up a bumper sticker, but in fact what he is talking about, to go into the littorals to get into the tough situation, to fight your way through and deliver power is exactly what we are doing [with DD 21].”27 More recently, big-ship admirals have begun deriding STREETFIGHTER vessels. As one admiral put it, “If the next major naval battle is fought in [Newport’s] Narragansett Bay, Streetfighters will be decisive.”28

Nor have the Development Command’s efforts influenced the Navy’s acquisition plans in any concrete way. Navy programs are currently dominated by incremental improvements to existing surface ships and aircraft. The service has yet to allocate any funds to procuring small, highly maneuverable ships such as STREETFIGHTER. Nor is that situation likely to change in the near future. In 2006, the Navy plans to begin building the CVX, a new aircraft carrier. It is therefore not surprising that the Defense Department’s top strategist has chided the Navy for failure to field experimental platforms.29

The U.S. Air Force and Unmanned Air Vehicles

The Air Force, a service historically defined by the technology of manned aircraft and dominated by fighter pilots, now faces the challenge of unmanned aerial vehicles, as well as military operations in space and cyberspace. In each case, the dominance of fighter pilots within the service has stymied innovation.

Rhetorically, at least, the Air Force sees itself in the vanguard of the RMA. As one recent article proclaimed triumphantly, “During the past decade, the U.S. Air Force has undergone a major transformation—a series of revolutionary changes so profound they have altered the face of modern warfare.”30 It has been a world leader in the development of stealth, precision-guided munitions, and the use of space to support military operations. As the official Air Force report on Operation ALLIED FORCE put it:

The air war over Serbia showed that the Air Force has embraced the RMA—not only in its acquisition strategies for emerging technologies, but in the way it used those technologies during this conflict. . . . The United States Air Force . . . showed that it is a leader in the revolution in military affairs by leveraging new concepts to support future joint and coalition efforts. . . . The air war over Serbia offered airmen a glimpse of the future, one in which political leaders turned quickly to the choice of aerospace power to secure the [Nato] Alliance’s security interests without resorting to more costly and hazardous alternatives that would have exposed more men and materiel to the ravages of war.31

Like the other services, the Air Force has begun to adapt conceptually and organizationally to the needs of the new security environment. It has reorganized itself into “expeditionary air forces” to project and sustain combat
power more efficiently. It has also developed the “Global Strike Task Force” concept, as a way of countering an adversary’s strategy for denying access to a combat theater. Along with the Navy, it is exploring such innovative concepts as “effects-based operations,” an idea that endeavors to link explicitly the application of military force to strategic objectives.

In fact, and notwithstanding its innovative concepts, the Air Force has as a whole been slow to embrace new ways of war. The hurdles it has faced in integrating unmanned airborne vehicles (UAVs) into its force posture are illustrative.

The service has, at least superficially, welcomed unmanned vehicles. It currently operates two squadrons of RQ-1A Predator medium-altitude-and-endurance UAVs. Controlled by ground-based operators, these aircraft transmit electro-optical, infrared, and synthetic-aperture-radar imagery via satellite to ground stations in the United States or the theater of operations.

It is also acquiring the RQ-4 Global Hawk, a high-altitude, long-endurance unmanned airborne vehicle designed to fly 12,500 nautical miles at an altitude of up to sixty-five thousand feet and remain aloft for thirty-six to forty-two hours. Advocates of the system argue that it is capable of replacing the venerable U-2 reconnaissance aircraft. The Air Force has formed a UAV Battle Lab to explore a number of novel operational concepts for the employment of unmanned vehicles. Perhaps more telling is the fact that in 1997 the Air Force awarded a UAV operator the Aerial Achievement Medal—roughly on a par in prestige with the Air Medal—for safely landing a damaged UAV at the Mostar air base in Bosnia-Herzegovina.

Last fall, the Air Force rolled out the first prototype “unmanned combat air vehicle” (UCAV), the X-45A. The aircraft, to be controlled by a ground-based operator, is designed to fly as high as forty thousand feet, have a thousand-mile range, and carry twelve miniature bombs. Its primary mission will be to attack enemy air-defense sites and pave the way for manned aircraft. The Air Force has also tested a weaponized version of the Predator as a rudimentary unmanned combat air vehicle.

Support for unmanned vehicles within the Air Force has, however, been lukewarm. The service’s modernization focus is upon a new generation of manned, short-range fighters to replace its existing ones; unmanned vehicles (and manned bombers as well) are being shortchanged. For comparison, the Air Force plans to spend nearly seventy billion dollars on the F-22 fighter aircraft and (along with the Navy and Marine Corps) at least two hundred billion more on the Joint

Because revolutions in military affairs disrupt long-standing norms and structures, it is not surprising that organizational resistance to change is one of the most formidable barriers to innovation.
Strike Fighter; the UCAV budget stands at a mere $126 million. In response to perceived foot-dragging on the part of the Air Force, Congress has passed legislation requiring that one-third of the nation’s deep-strike capability be unmanned by 2010.

The cultural barriers against embracing unmanned vehicles are substantial. UAVs have been in use for decades, but the Air Force has yet to exploit them fully. Over the past two decades, the Defense Department has spent two billion dollars on unmanned airborne vehicles—roughly the cost of a single B-2 bomber, one-tenth the money it spends on manned combat aircraft in a single year. As a result, UAV technology remains far short of its potential. Indeed, in 1993 Congress created the Defense Airborne Reconnaissance Office to manage unmanned-vehicle programs after unsuccessfully prodding the Pentagon to take them more seriously. The Air Force formed its UAV squadrons only after the Army threatened to take the mission—and the associated resources—away from it.

The pilot culture that dominates the Air Force is another obstacle. While Air Force UAV operators must be pilots, tours with UAV squadrons are designated as nonflying assignments and are thus less than desirable. As an incentive for serving two years with a Predator squadron, the Air Force has been obliged to give pilots the subsequent opportunity to fly a new type of aircraft, which would improve their career chances.

The emergence of UAVs and UCAVs has created growing tension between pilots and supporters of unmanned systems. Many pilots see the UCAV as a threat. As one officer put it, no one “has ever succeeded in picking up a woman in a bar by saying he commanded a wing of drones.” While humorous, such sentiment illustrates the barriers to adopting new approaches to combat. This situation is analogous to that in the 1950s, when the advent of intercontinental ballistic missiles threatened the manned-bomber community.

WHAT IS TO BE DONE?
The services have so far failed to match the rhetoric of transformation with action. While each claims to embrace new ways of war, none has yet demonstrated a sustained commitment to fundamental change. Nothing shows this more clearly than their acquisition budgets. Service funding is still dominated by incremental improvements to traditional systems; radically new technology, doctrine, and organizations have received smaller resources. None of this should be surprising. Large bureaucracies such as the U.S. armed forces are designed to minimize uncertainty, including that brought on by large-scale change. And new is not always better. Yet the U.S. armed forces face the imperative of adapting to the new and different challenges the United States will face in coming
years. Should they fail to do so, they could find themselves becoming increasingly irrelevant.

It would be wrong to view the services as uniformly opposed to fundamental change. Rather, each service is split between traditionalists and elements who are enthusiastic about new ways of war. One recent survey of the U.S. officer corps revealed significant splits over the character and conduct of future wars as well as over the urgency of change. The Defense Department needs to identify and nurture forward-looking constituencies. The starting point should be an intellectual map of the services, one that identifies and locates both support for and opposition to new mission areas. Such a map could assist the Defense Department’s leadership in channeling resources to those portions of the services that are most enthusiastic about emerging warfare areas. It could also assist the department in evaluating the adequacy of military career paths.

The Defense Department also needs to devote additional resources to experimentation. In particular, the services should advance from the stage of war-gaming innovative concepts to acquiring small numbers of the weapon systems involved and developing concepts and organizations for their use. The Navy, for example, should purchase a squadron of STREETFIGHTERS to form an operational test bed for network-centric warfare. The Marines, for their part, should establish experimental units dedicated to projecting power in the face of capable access-denial defenses and to conducting military operations in urban terrain.

More generally, the Defense Department should begin redistributing resources away from legacy systems of declining utility and toward new ways of war. The Pentagon should scale back or cancel weapons that are heavy or have limited mobility, highly detectable signatures, and limited range; it should increase funding for long-range precision strike, stealth, and C4ISR* systems. The department should also increase substantially the funds it devotes to research and development.

Today’s defense budget is split fairly equally between the services. While such an arrangement minimizes interservice friction, it is not particularly conducive to innovation. Indeed, there is a strong argument to be made that interservice competition can be an engine of change. One way to promote innovation would be to force the services to compete for funds based upon their ability to meet current and anticipated operational and strategic challenges. These challenges

* Command, control, communications, computers, intelligence, surveillance, and reconnaissance.
would include the need to assure access to regions of critical importance to the United States; gain and maintain information and space superiority; protect against nuclear, biological, chemical, and information attack; and conduct military operations in urban terrain. In order to ensure that the American armed forces meet these emerging challenges, the secretary of defense should set aside a significant portion of the military’s procurement budget for innovative programs.

The service secretaries are a potentially powerful but generally underutilized constituency for change. They have it within their power—through control of promotion boards and officer assignments—to have enduring impacts on their services. They should wield this power to ensure that officers associated with emerging warfare areas, such as space and information warfare, enjoy opportunities to rise to senior leadership positions.

The United States leads the world in many of the technologies that are driving the information revolution, as well as many of the weapons that the revolution has spawned. Transforming the armed forces will require the Defense Department not only to continue to acquire advanced weapons but to develop the organizations and doctrine needed to employ them effectively. That attempts to do so have encountered resistance is not surprising. Change is by definition a disruptive process, one that creates winners and losers. Still, the U.S. armed forces must change radically—adding new capabilities and shedding old ones—if they are to meet the challenges of the emerging security environment.

NOTES

8. James R. FitzSimonds [Capt., USN], “The Cultural Challenge of Information


20. Ibid., p. 28.


27. “Murphy Slams ‘Street Fighter,’ Navy Distances Itself from Comments,” Inside the Navy, 18 October 1999, p. 3.


NEW INSIGHTS FROM OLD BOOKS

The Case of Alfred Thayer Mahan

Jon Sumida

Alfred Thayer Mahan’s *The Influence of Sea Power upon History, 1660–1783*, which appeared in 1890, is widely regarded as the first important study of the relationship between naval affairs and international politics. Mahan subsequently published twenty-odd additional volumes that extended and elaborated the views presented in this book. On the present occasion, an article based upon the traditional summary of Mahan’s main ideas could be justified as an obligatory nod to the U.S. Navy’s intellectual heritage, or as an act of faith in the capacity of patristic writing to inspire strategic insight. Recent scholarship, however, has demonstrated that Mahan’s thinking about sea power has been fundamentally misunderstood. This article will thus examine three areas where the new interpretation of Mahan affects consideration of problems that are of interest today. The first is naval and military cooperation when fighting in inland or coastal waters; the second is the nature and role of naval supremacy with respect to a complex world system of trade; and the third arises from the requirements of higher naval education in a period of rapid technological change. In other words, Mahan’s work will be related to jointness and power projection from the sea, the expansion of the global economy, and the “revolution in naval affairs.”

There are three main arguments. First, Mahan believed that when one side in a conflict possessed
absolute sea command or in special cases even temporary local control, naval operations in direct support of land forces could be of decisive importance. Second, Mahan maintained that naval supremacy in the twentieth century would be exercised by a transnational consortium acting in defense of a multinational system of free trade. Third, Mahan was convinced that the transformation of naval materiel by radical technological change had not eliminated tactical and strategic uncertainty from the conduct of war, and that the improvement of executive ability through the rigorous study of history should therefore be the basis of naval officer education.

Mahan is often portrayed—because of misreadings of fragments of his writing, or all too often upon no reading of the original texts at all—as a purveyor of truisms about naval strategy and doctrine. The resulting caricature is frequently either misapplied or dismissed as outdated. This article, which is informed by the study of all of Mahan’s major publications and surviving correspondence, should remind us of the merits of the adage, “When you want a good new idea, read an old book.”

A COMPLEX PICTURE OF THE INTERRELATED DYNAMICS

Alfred Thayer Mahan was an officer in the Union navy during the Civil War. He was never a participant in a major battle, but his active service included many months of inshore work in small warships enforcing the blockade of the Confederate coast. Nearly two decades after the end of hostilities, Mahan accepted a commission to write a book about naval operations on the Caribbean coast and up the Mississippi and Red Rivers in the war. In addition to drawing upon his own experience during this conflict, Mahan studied memoirs and documents, and corresponded with veterans from both sides. The completed work, which was entitled The Gulf and Inland Waters, was published in 1883. Several years after the appearance of Influence of Sea Power upon History and its two-volume sequel, The Influence of Sea Power upon the French Revolution and Empire, which came out in 1892, Mahan produced a biography of the admiral who had commanded most of the Union operations described in his first book. Admiral Farragut, which was published in 1897, gave Mahan another opportunity to present his views on fighting in littoral and interior waters that involved cooperation between the army and navy.

During the American Civil War, the lack of a fleet meant that the Confederacy could not mount an effective challenge to Union control of the high seas. Moreover, the naval weakness of the southern states exposed their vital internal riverine communications and major ports to seaborne assault. Over the course of the four-year conflict, the territorial integrity and economic vitality of the South were compromised by the integrated action of the Union army and navy,
which established Federal control of the Mississippi and captured New Orleans and Mobile. Mahan’s two accounts of these campaigns demonstrate that he possessed considerable knowledge of the special characteristics of “brown water” fighting, appreciated the necessity of connecting the activity of land and naval forces, and recognized that the success of joint operations had been a major contributor to the ultimate Union victory.

In books written before and after the Farragut biography, Mahan criticized Admiral Horatio Nelson’s advocacy of amphibious operations in support of land campaigns and, in general, opposed overseas expeditions. But these views were applied to circumstances in which the opposing side possessed—or was supposed to possess—the capacity to dispute sea command. Mahan reasoned that in such a case any attempt to project power from water to land risked naval assets that were needed to preserve the general control of the oceans, upon which all depended. When the maintenance of maritime lines of communication was not an issue, he had no objection to using naval force in combination with an army to achieve a military objective, and he well understood that such action could have great strategic value.

Indeed, Mahan attributed his initial inspiration—for the idea that naval supremacy was of much larger historical significance than was generally recognized—to his reflections on a historical case involving the use of uncontested command of the sea to achieve decisive military success. In his memoirs, he recalled that in 1885 he had chanced upon Theodor Mommsen’s history of ancient Rome. While reading this book, Mahan had been struck by the thought that the outcome of the wars between Rome and Carthage would have been different had the latter possessed the ability, as did the former, of using the sea as an avenue of invasion instead of moving its armies over land.

After some reflection, Mahan decided to apply the example of the victory of a state that could use naval force effectively over one that could not to the history of European wars in the late seventeenth and eighteenth centuries. This resulted in the first of the “influence of sea power” volumes, in which Mahan closed the introduction with a lengthy examination of the naval aspects of Rome’s defeat of Carthage. He ended the main narrative of The Influence of Sea Power upon History with an account of the British defeat at Yorktown in 1781. The outcome of this battle had been determined by the reinforcement of American and French armies by sea, and also by French naval control of surrounding waters, which had prevented a British fleet from relieving the besieged British army. The Yorktown disaster had prompted negotiations that ultimately ended the war and

When dealing with Mahan, the focus of inquiry should not be his statement of principle but his choice of issues, and the complexities of the historical cases that were his main subjects.
established American independence. In the book that made his reputation, Mahan thus used the survival of what was to become imperial Rome and the creation of the United States as powerful historical testaments to the transcendent value of naval force in support of military operations.

But *The Influence of Sea Power upon History* also introduced a set of propositions about the relationship between the economic basis of national strength and the development and effective use of a navy. Seaborne trade, Mahan maintained in his first best-seller, was a critically important generator of wealth. In the event of war, a nation that could protect its own maritime commerce while disrupting that of its opponent could shift the balance of national resources decisively in its own favor. A fleet capable of winning and keeping command of the sea was required to accomplish both of these tasks. In peace, therefore, a great state was well advised to do everything it could to build the strongest possible navy. Over time, the cumulative effect of sound naval policy and strategy in peace and war was economic prosperity and territorial aggrandizement.

Naval force structure and deployment were also important variables. Cruiser attacks on scattered shipping, Mahan believed, were incapable of inflicting prohibitive losses on a large merchant marine. Blockade of the enemy’s main ports—implemented by a fleet of battleships capable of defeating any force that was sent against it—was the only way to accomplish the complete or near complete stoppage of overseas commerce required to achieve a significant strategic effect against a great maritime power. It was for this reason that Mahan made the number of battleships the measure of naval potency, and the destruction of the enemy battle fleet through decisive engagement—for the purposes of either securing or breaking a blockade—the main operational objective of naval strategy.

These interrelated arguments addressed major concerns of Mahan’s own time. From the 1880s, the general expansion of European navies in response to increasing imperial rivalry was accompanied by intensive debate over the relative merits of a naval strategy based on commerce raiding by cruisers, as opposed to one based on command of the sea by battleships. In addition, the advent of steam propulsion and metal hulls had vastly increased the efficiency of maritime transport, which in turn caused a sharp upturn in overseas commerce and the wealth generated by this kind of activity. Mahan’s choice of European great power conflict during the late age of sail as the vehicle for his argument also favored discussion of the general struggle for naval supremacy in preference to case studies of combined operations along coasts and rivers. So although Mahan clearly recognized the importance of power projection from sea to land, it was his examination of the contest for command of the sea, and its political-economic consequences, that created the immediate and wide audience for *The Influence of Sea Power upon History* and later publications. The resulting
association of Mahan with arguments exclusively about naval supremacy distorted perceptions of his identity as a strategic theorist, setting the stage for misleading comparisons with writers (such as C. E. Callwell and Julian Corbett) who focused more on the relationship of land and sea power. But a far greater problem was created by the serious misunderstanding of the basic character of Mahan’s rendition of European naval history in the age of sail, a misperception that led to faulty inferences about Mahan’s fundamental views on grand strategy.

The “influence of sea power” series began in the mid-seventeenth century with a situation in which three major maritime states—France, the Netherlands, and England—were roughly balanced with respect to naval prowess and accomplishment. It ended in the early nineteenth century with the wars of the French Revolution and Empire, during which Britain’s Royal Navy more or less ruled the waves. In addition to the two works named previously, which provided an overview of the entire period, Mahan wrote two supporting case studies: a biography of Admiral Nelson, and an account of the War of 1812. In terms of “plot,” the entire series could be read as the story of the rise of Britain’s naval supremacy and its consequent achievement of economic and political preeminence in Europe. In terms of “moral,” the series seemed to say that Britain’s sustained and aggressive use of a large fleet to obtain territory, wealth, and power could be emulated by any state that had the mind and will to follow its example.

Mahan, it appeared to many, had intended his analytical history to be a grand strategic primer for his own times, and in particular for the government of his own country. He was indeed a proponent of a much strengthened U.S. Navy. It is thus not hard to imagine that he hoped that his homeland would become the world’s greatest power in the twentieth century by the same means that Britain had used to achieve this status in the period covered by his histories. The fact that the United States ultimately rose to the top in large part through the effective use of naval supremacy has only reinforced the propensity to draw such inferences about Mahan’s underlying motive.

But careful consideration of Mahan’s actual writing in the “influence of sea power” series, of his political-economic outlook, and of his punditry about the future course of world politics makes it impossible to accept the foregoing characterizations of his account of naval warfare in the late age of sail and of its intended application to the twentieth century. The first installment of the “influence of sea power” series is about the failure of France to exploit its maritime assets properly, a failure that in Mahan’s view allowed Britain to achieve major...
successes in war virtually by default. Mahan chose to close the book with a disproportioneately lengthy account of the American Revolution, a conflict in which sound French policy and deployments resulted in Britain's defeat and the loss of a vast and rich colonial territory. In the wars of the French Revolution and Empire, in contrast, the navy of France was compromised from the start by political upheaval and institutional disintegration. The second installment was thus about Britain’s use of naval supremacy to contain a militarily preeminent France through a strategy of attrition. Mahan did not hold that the ultimate outcome had been preordained—that is, that naval supremacy as such guaranteed victory. The triumph of Britain, given the evenness of the balance between the opposing sides, he argued in both the second and the third installments, depended upon extraordinary operational naval leadership—in the person of Nelson. In the concluding fourth installment, Mahan’s main theme was that inadequate American naval strength was the fundamental explanation of diplomatic failure before the War of 1812 and of naval operational impotence, with all its attendant serious strategic drawbacks, during the conflict.

Britain and British naval strategy did not, in short, represent the focus of the “influence of sea power” series. Mahan’s histories did not constitute a simple morality play about a single state acting according to a prescribed general course of action; they offered instead a complex picture of the interrelated dynamics of naval and maritime commercial activity on the one hand, and international politics on the other. Mahan’s essentially liberal political-economic views, moreover, led him to reject the mercantilist conception of a world consisting of competing players with mutually exclusive interests. Mahan believed that free trade between nations promoted increases in the volume of international exchanges of goods, which worked to the benefit of all participants. The great expansion of French overseas shipping after the War of the Spanish Succession, he argued in the first installment of the “influence of sea power” series, was attributable to peace and the removal of restrictions on commerce, not to government initiatives. In the second installment, Mahan observed that sea power was an organism that included not only organized naval force but free maritime enterprise. While the former depended upon state funding and direction, the latter thrived in the absence of government interference. During the wars of the French Revolution and Empire, Mahan maintained, the British state was able to exploit the prosperity produced by an international sea-based mercantile system that it could protect but did not possess. Britain was not, in other words, the owner of sea power but its custodian.

Mahan believed that Britain had been both the defender and main beneficiary of seaborne trade in the late eighteenth and early nineteenth centuries because its parliament had been dominated by a small group of men with close
ties to maritime commerce. Such an oligarchy had been predisposed to heavy spending on the navy, producing a fleet strong enough to defend a merchant marine that carried a large proportion of the world’s overseas trade. Over the course of the nineteenth century, however, the democratization of the British political system undercut the manipulation of government policy by a mercantile elite. As a consequence, Mahan argued, the British state of the late nineteenth and twentieth century lost the will to finance a navy capable of defending what had become a much larger and increasingly multinational system of oceanic economic exchange. Moreover, in Mahan’s view, no single democratized power would be capable of assuming such a burden. For this reason—and because he was convinced that free trade conditions provided large benefits to all major maritime countries—Mahan concluded that in the twentieth century, naval supremacy would be exercised by a transnational consortium of navies. The basis of such a system, he insisted, would not be formal agreement but the absence of important conflicts of political interest, coupled to a common stake in the security of a highly productive form of economic activity. Mahan was thus convinced that Britain and the United States would cooperate without recourse to a treaty and that in such a relationship the latter would serve as the junior partner. To play even this supporting role effectively, Mahan insisted, America needed a larger navy. He did not advocate the creation of an American navy that was stronger than every other unless the British navy was weakened by inadequate financing or by war with a European competitor.

Mahan offered his views on the future course of international affairs in several book-length monographs and in periodical articles that were later collected and published as books. In them Mahan contemplated a range of possible courses of events. These included the containment of an expansionist Russia by an international coalition, war between Britain and Germany, and even a cataclysmic collision between European and Asian civilizations. What he did not do was apply a crude reading of the great-power contests of the late age of sail to the industrial future and thereby imagine the rise of a hegemonic United States through offensive naval warfare and mercantilist economic policy. While his realist temperament prompted him to argue that war and the threat of war were likely to be facts of life for the foreseeable future, Mahan did not rule out either the possibility or desirability of general peace founded upon the workings of an international system of free trade. In such a world economy, he was confident,
the energy and entrepreneurial spirit of the American people would enable them to compete successfully.

In the second half of the nineteenth century, the onset of industrialization transformed naval materiel within the span of a generation. When Mahan was a midshipman at the U.S. Naval Academy, just before the American Civil War, he was trained on wooden sailing ships armed with muzzle-loading guns. By the time he retired from the service at the end of the century, steel warships propelled by steam and equipped with breech-loading weapons of much larger size and power had become standard. The sudden obsolescence, as a result of rapid technical change, of much of what had constituted traditional naval fighting practice and the virtually worldwide sense that what now really mattered in war was the possession of the latest, and therefore most capable, naval armaments undermined the self-confidence of naval executive leaders. Conversely, naval officer technicians could celebrate the wonders of technical improvement and claim that the critical importance of qualitative advantage in materiel made their activity central to the efficiency of the Navy. Moreover, administrative burdens had been magnified by the management needs of the new technology and also by the expansion of the American fleet that had begun in the 1880s; these factors created a large class of naval officer bureaucrats with pretensions to higher status not directly connected to traditional requirements for command at sea.

The relative decline of naval officer executives alarmed Mahan. By dint of intellectual patrimony and personal experience in the greatest conflict ever fought by his service up to his time, he had decided opinions on the paramount value of effective leadership in war and how it might be developed. Mahan’s father, Dennis Hart Mahan, a distinguished professor at the U.S. Military Academy at West Point, had believed that great executive leadership was of crucial importance in war. The elder Mahan had observed that at critical junctures, commanders would be confronted with complex, contingent, changing, and contradictory information, which meant that decision making could never be reduced to the mechanistic application of rules or principles. Development of the temperament required for sound judgment under such circumstances, Dennis Hart Mahan was convinced, could be aided by the study of detailed and analytically rigorous operational history. There can be little doubt that this outlook was imparted to his son, and thereafter reinforced by the younger Mahan’s direct observation of command decision making in the Civil War. Alfred Thayer Mahan’s first publication, in 1879, was an essay on naval education in which he attacked what he regarded as the overemphasis on technical subjects and called for much greater attention to the study of what amounted to the liberal arts. Such an approach, he maintained, would develop the moral qualities that officers required
to be able to make decisions in the face of danger and uncertainty. The vital role of moral strength with respect to executive command, along with the appropriate means of improving it in naval officers, became a theme in Mahan’s later writing that was no less important to him than his examination of the relationship between naval affairs and international politics.

In *The Influence of Sea Power upon History*, Mahan argued that although tactics changed as the character of armaments altered, the validity of the basic principles of strategy were relatively unaffected by technical progress, and human character was an absolute, a constant. History, therefore, might have little to say of current applicability to tactics but a great deal that was pertinent to strategy and operational command. Mahan devoted as much attention in the main narrative of this work to the strategic direction of naval operations as he did to his grand strategic argument about the relationship between naval supremacy and the course of international politics. He also made a few observations about the critical effect of individual moral character on the exercise of naval command. In later installments of the “influence of sea power” series, he remained no less attentive to strategic questions and, through his treatment of Nelson’s leadership qualities, wrote at length about the moral dimensions of executive decision making in war.

In several of his articles, Mahan maintained that the essence of effective command comprised rapid and judicious risk taking and full responsibility for outcomes. This set of characteristics was alien to the scientific *mentalité* of the engineer, who dealt deliberately with the discovery of certainty about physical matters through controlled experiment, and to the bureaucratized mindset of the administrator, who countenanced delay and fragmented accountability. In peace, an executive leader had few if any opportunities either to display his capacity for war command or to acquire experience that would enable him to develop it, while technicians and bureaucrats flourished in the pursuit of engineering innovation or administrative expansion. For Mahan, therefore, serious naval history, of the kind that he had produced in the “influence of sea power” series, served two major practical functions. First, it reminded the navy of what executive war command was and why it was important; second, it provided a sound educational basis for developing that capacity in officers who had no war experience. The latter task was accomplished through stories about naval decision making in war, narratives that prompted readers to imagine the psychological dynamics as well as material circumstances that condition the direction of operations in real conflicts.
Mahan lacked the powers of technical ratiocination that were needed to evaluate properly a complex engineering problem, such as capital-ship design. His criticisms in the early twentieth century of the all-big-gun battleship, therefore, failed to take into account several significant factors that exposed his analysis to swift and thorough destruction. But neither was Mahan a naval technological Luddite. If he was a critic of many of the claims made for mechanical innovation, it was because he was convinced that such progress had not eliminated uncertainty from decision making in war, and that the decadence of the naval executive ethos that had resulted was thus a dangerous weakness. His antidote to the technological determinists of his time, however, was history rather than political science; he believed that the verisimilitude afforded by detailed narrative about things that had actually happened could engage the minds and feelings of students of command in ways that summarized lessons or abstractions could not. Mahan's preference for historical representation over the construction of explanatory systems when dealing with the past is in line with much that has been argued recently by proponents of chaos and complexity theory. Further, his remedy for moral dilemmas—confidence in intelligent intuition—is one that is supported today by the findings of cognitive science. Viewed in light of these modern, cutting-edge inquiries into human learning and behavior, the writings of Mahan may be regarded as not just relevant but revelatory.

A COGNITIVE POINT OF DEPARTURE

For nearly a hundred years, Alfred Thayer Mahan's pronouncements on naval affairs and international politics were too famous to be ignored but also too extensive, difficult, and complicated to be easily understood as a whole. From the start, most writers on naval history and strategy misperceived his work, and successive generations compounded the errors of their predecessors, creating a large literature whose shortcomings further obstructed access to the meaning of the original texts. As a consequence, Mahan's basic ideas have been misrepresented as follows: first, sea control was always the central question of naval strategy; second, the ideal of national grand strategy was the achievement of naval supremacy as the prerequisite to international economic and political preeminence; and third, success in naval warfare depended upon the correct application of certain principles of strategy. These propositions add little to current naval discussions, which consider the American possession of sea control and a monopoly on superpower status virtually as givens, and which are preoccupied by the transformation of fighting practice through radical technological innovation.

The major arguments of Mahan that can be found through comprehensive and rigorous critical examination, however, are very different from what has
been supposed. Moreover, the issues that prompted him to put pen to paper were remarkably similar to those of today. He began both his naval and writing careers dealing with joint operations in coastal waters. Mahan was confronted by the rapid expansion of a global system of free trade; by uncertainty about what America’s proper naval role under such conditions should be; and by a “revolution in naval affairs” that was occasioned by the replacement of pre-industrial by industrial naval armaments and that in turn raised large questions about the nature of wartime command and the education of those who would exercise it.

Mahan’s contemplation of these problems produced the following conclusions. First, close cooperation between land and sea forces is essential for the success of joint operations, whose outcomes can determine the victor in a major war. Second, because the cost of building and maintaining a navy that is strong enough to command the seas unilaterally will be too high for any single power, sea control in the twentieth century and beyond would be the responsibility of a transnational consortium of navies. Third, great advances in technology do not diminish reliance upon the good judgment of naval executive leaders, who could best be prepared for effective high-level decision making in war by the proper study of history.

Identifying Mahan’s true basic attitudes toward power projection from sea to land, naval supremacy, and the relationship between technological change and naval officer education does more than correct academic error. What have been believed to be Mahan’s ideas created a body of theory that still—whether through acceptance, modification, or rejection—forms part of the thought processes of most senior naval officers. Changing what has for so long been a cognitive point of departure, therefore, has significant implications for anyone concerned with the future of naval policy. Mahan has been widely regarded as the discoverer of what he supposedly believed were universal truths about naval strategy that were to be applied directly. The fact is, however, that Mahan’s propositions were observations about particular phenomena rather than general lessons. When dealing with Mahan, the focus of inquiry should therefore not be his statement of principle or delineation of precedent but his choice of issues, and the complexities of the historical cases that were his main subjects. The crucial linkages between his past and our present, in other words, are not to be found in his conclusions but in his questions and his conduct of the inquiry. These are still worth engaging, because Mahan faced problems that were similar to those that confront navies today, and he brought to them a powerful intelligence that was informed by rich experience and wide reading. History was the venue for Mahan’s scholarly labors, because he understood both the limits of theory and the power of narrative when it came to matters of human behavior and social
organization under the conditions of war. There is much more that can and
should be written about the general and particular aspects of navies and naval
power, but approaching, let alone matching, the intellectual standard of Mahan's
pioneering achievement will not be easy.
THE COAST GUARD AND NAVY

It’s Time for a “National Fleet”

Colin S. Gray

The U.S. Coast Guard and the U.S. Navy have been, are, and will remain complementary. They are not competitors. However, notwithstanding the distinctiveness of their missions and functions, in practice their duties overlap. There is a zone of activity wherein the services share maritime geography and foci of concern. Because the Coast Guard operates shallow-draft warships, it can be misrepresented as the coastal or shallow-water navy of the United States. Similarly, because the Navy supports the Coast Guard when necessary and feasible, perception of naval enthusiasm for such support (and beyond) can feed ill-founded anxiety that the Coast Guard is in peril of imperial absorption by the much larger service. A well-ordered U.S. defense community, confident in its understanding of the emerging strategic environment and prepared to pay the freight for national security, would provide little fuel for these essentially foolish apprehensions. However, this article is propelled by the appreciation that even though the Coast Guard and the Navy are natural and necessary allies, trends exist today—both internal to each service and, even more, in their contexts of operation—that could strain their relationship.

As we shall see, it is not surprising that most of the sources of difficulty in the interservice relationship stem from questions about missions and equipment pertaining to the Navy, rather than to the Coast Guard. The latter does not face challenges to its roles, missions, and relative importance that are so radical as those that stalk the Navy. The Coast...
Guard, understandably, is occasionally anxious lest some of its duties be outsourced, privatized, or picked up by a Navy looking for self-justifying tasks. However, those periodic perils (real and imaginary) fade nearly into insignificance when compared with the vulnerability of the Navy to shifts in defense-intellectual fashion and foreign-policy mood. The Coast Guard’s potential (domestic) critics are largely toothless tigers; the Navy’s are not. The centerpiece of this discussion, then, is the future relationship between the Navy and Coast Guard in light of their common status as sea services of the United States, under the conceptual umbrella of a “national fleet.”

That relationship cannot be considered in isolation, however. Both services must shape their interconnection with reference to powerful contextual factors. Whether or not Navy–Coast Guard relations constitute a love match, each needs the other. Trends point with a uniform logic to the common sense contained in the idea of a national fleet. What the Navy lacks by way of blue-water challenge from a pressing “high end” threat finds ample compensation in opportunities and problems posed by the emerging information-led revolution in warfare. An unmistakable trend afflictng the all-high-end U.S. Navy is a declining number of ships. Fleet size is not everything, but—as the last Chief of Naval Operations reminded us—“numbers do matter.” Especially do numbers count when operations of all kinds must be conducted worldwide by a rotational deployment pattern.

Just as the Navy’s operational tempo has become unsustainably high for a peacetime rotational fleet, so the Coast Guard is obliged to cope with a higher demand for its services. The uses made of the sea, which taken all together constitute the principal driver of Coast Guard activity, have risen, are rising, and are projected to rise much farther yet. Quite aside from its national defense mission, the Coast Guard has a basket of traditional duties, a collection expanding in variety, quantity, and quality of challenge.

THE ARGUMENT
We will examine in some detail the current and anticipated conditions and circumstances of the Navy and the Coast Guard, and also the terms by which and ways in which they can best complement each other. First, however, it is useful to break the rules of dramatic construction and reveal the five points that, together, represent the “argument” of this article. These points are not particularly remarkable or even controversial (except for the fifth). Rather, as we will see after setting the argument, actual and potential controversy on a major scale attaches to the character of the “contexts” (political, military, intellectual-doctrinal, societal, and so forth) within which the services must operate over the decade ahead.
For the Navy in particular, there is much scope for dispute as to the requirements in military and strategic effect that U.S. foreign policy will place upon it.

First: The U.S. Navy needs to be all that it has to be as a war-fighting instrument. Indeed, it is as the exemplar of naval prowess that the Navy supports the foreign policy of the global superpower. Given that major naval combatants, with midlife refits, may well be in commission for thirty or forty years, contemporary defense-intellectual fashions and fads, even today’s “best guesses” about the strategic future, should be treated with some reserve. Much as nominal, back-of-the-envelope answers to, say, ballistic-missile-defense queries can always be manufactured promptly on demand (how long does it take to produce a vugraph?), so high-end naval power—in general and specifically—can quite easily be dismissed as yesterday’s unaffordable and irrelevant answer to the bold novelties of tomorrow. The past century has seen great combat fleets “sunk” virtually by torpedo-firing flotilla craft, erased by airpower, sidelinied or destroyed by nuclear menace, and now relegated to deep-reserve status by “asymmetric” anti-access capabilities that could, allegedly, lock out most of the U.S. Navy from an enemy’s maritime approaches.

Let us note that in the twentieth century great navies survived strategic challenge from new classes of threats deriving from the subsurface, aerial, space, and electronic environments. The smart money says that the U.S. Navy will be as successful in the future as it was in the past at riding the erratic waves of revolutions in military affairs, and that predictions of its imminent strategic relegation will prove as ill judged in the future as before. Because “history is geography,” as Jacob Burckhardt wrote with only modest exaggeration, the U.S. Navy cannot responsibly be regarded as just one player on an infinitely fungible national security team.

Second: Insofar as practicable, the Navy should leave coast guarding to the Coast Guard. If naval warships happen to be in the right place at the right time and have no truly urgent and stressful national-defense missions to perform, then by all means let them contribute to security at sea in its fullest and broadest sense. However, just as the warships with orange and blue stripes can fight but are not expected to excel in the conduct of naval warfare, so the gray hulls of the Navy certainly can function as surrogate coast guard but are not expert in that role. Moreover, as the Navy continues its up-market movement with acquisition of supercapable and superexpensive destroyers and cruisers, so it will have ever
fewer (leading to zero) relatively small and cheap general-purpose warships (that is, frigates) appropriate for shallow-water duties short of naval war.

Given the authoritative status of the national fleet “treaty” of 21 September 1998 between the Navy and the Coast Guard, it is troubling to find that the *Navy Strategic Planning Guidance with Long Range Planning Objectives* of April 2000 is all but oblivious both of the Coast Guard’s skills and of the limitations of the Navy itself. The *NSPG* sees no tension between arguing soundly that “insufficient numbers entail strategic risk as well as excessive personnel and operational tempos” and reaching close to coast-guarding domain. Describing the “day-to-day engagement” activities of the Navy, it argues that

this engagement process also encompasses the spectrum of military operations other than war (MOOTW), which repeatedly employs naval forces in missions such as humanitarian disaster relief, non-combatant evacuation operations (NEO), peace support missions, enforcement of embargoes and no-fly zones, counterproliferation measures, and rapid reaction to terrorism. Future forces also must be prepared to support law enforcement agencies to deal effectively with non-military challenges to our national security, such as illegal immigration, illegal drug trafficking, and other international criminal activity.

There are times when what U.S. policy most needs for its support is what Wayne P. Hughes, Jr., usefully terms “formidability—e.g., as manifest bigness, obvious speed, devilish appearance, and demonstrated cunning.” However, as a general rule, sending an *Arleigh Burke*–class guided-missile destroyer like the USS *Cole* into the Arabian littoral is akin to “sending a man out to do a boy’s job.” A U.S. Navy properly and jointly directed in strategic terms should focus on the higher-risk tasks that fit logically and prudently with its trends in equipment acquisition, leaving as many extramilitary maritime security tasks as possible to the Coast Guard. Of course the Navy could undertake many coast-guarding jobs, and in many countries navies do just that (Britain’s Royal Navy, for example). However, that has not been the American way, and given the fact that the United States has the world’s best coast guard (as well as the world’s best navy, by far), there is every reason for exploiting the duality of its sea services.

*Third: The national fleet treaty of 21 September 1998 is no more than common sense.* The “Treasury Fleet,” later the Coast Guard, has served in or with the U.S. Navy in all of America’s wars that have had maritime dimensions. The concept of a national fleet can raise paranoid fears in both services, but in fact it dignifies what has been the American experience: the Navy and Coast Guard have fought as a national fleet. The Navy’s anxiety that the Coast Guard would use the national-fleet idea as a license to hunt for more missions, more force structure,
and more money at the Navy’s expense is amply matched by Coast Guard nervousness lest the national fleet should become a “bear hug,” extinguishing organizational independence, even autonomy. Neither fear is well founded, but both need exposure and recognition.

The fact of the matter is that on its current course regarding technical choices and force structure, the U.S. Navy is opening more and more space between the shallow-draft multimission warships that major Coast Guard cutters are and its own newly “low end” destroyers (DDs) and guided-missile destroyers (DDGs). Should the Navy, improbably, elect to pursue variants of the STREETFIGHTER concept, then indeed it would be back in the business of small (and more expendable) warships. At present, though, that point is strictly speculative. It is no stretch to reason that a navy acquiring DDG 51s (Arleigh Burkes) and intending (and hoping) to buy DD 21s as the low end of its fleet mix is likely to be size challenged. The concept, and one hopes the practice, of the national fleet provides an important, actually essential, part of the answer to the challenge of too few “frigates.”

Fourth: The Coast Guard should be recapitalized by rapid implementation of its planned Integrated Deepwater System (IDS). For an acquisition cost in the range of fifteen to twenty billion dollars over twenty years, the Coast Guard will replace and modernize with a new class of maritime security cutter, long-range aircraft, and C4ISR.* This modernization is necessary to enable the service to cope with the expanding demands of its maritime safety, environmental protection, and law enforcement roles, as well as its national defense duties.

Coast Guard maritime security cutters will not be frigates, but these categories of warship do bear more than a casual similarity. Writing at the end of the Cold War, Eric Grove informed us that a frigate is a “combatant of about 1,750 to 3,000 tons usually optimized for ASW [antisubmarine warfare] but with general purpose capability; essentially intended for the escort of noncombatant shipping, although useful for patrol and limited offensive operations, capable of oceangoing transits and tasks; usually air capable [operating helicopters].”

Oliver Hazard Perry–class guided-missile frigates (FFGs) displace around four thousand tons and the Coast Guard’s Hamilton–class high-endurance cutters (WHECs) about three thousand. That is to say, some Coast Guard cutters are frigate-sized warships. They can be made as frigate-like as the Coast Guard prefers, politically is permitted, and can be afforded. Looking to the future, however, when the Navy ceases to have FFGs in commission, no one will be in any

---

* Command, control, communications, computers, intelligence, surveillance, and reconnaissance.
danger whatsoever of mistaking a hull with an orange stripe—even one from the top end of the catalog—for a nine-thousand-ton DDG.

Service sentiment aside, the national fleet of the United States needs some more or less frigate-sized and modestly priced general-purpose warships. If the Navy cannot afford to operate them, given the military logic of its high-end-procurement strategy, it should be a cause for rejoicing that the Coast Guard, prudently modernized, can take up some of the slack. If a modernized Coast Guard is to complement the Navy in a national fleet—or simply in joint warfare—it requires equipment that not only is complementary but can operate seamlessly with that of the other armed services. Whether or not some in the Navy regard the Coast Guard as a necessary evil (or regrettable necessity!), all can agree that a modern Coast Guard, if it is to sail and fly on national-defense duties, must not be a net liability to operational and tactical commanders. A Coast Guard recapitalized by the IDS has to be able to function with the Navy as a tactically and technically reliable team player, working by the same rules of network-centricity.\footnote{Needless to say, it comes down to money, doctrine, planning, and training.}

Fifth: As an armed service, the Coast Guard probably should relocate to the Department of Defense. There is no general law governing the proper character or bureaucratic location of a coast guard. Each country does it its own way. Moreover, what arrangement is suitable may vary, depending upon trends in roles and missions of both the service in question and host institutions. The U.S. Coast Guard, consolidated as such in 1917, lived under the aegis of the Treasury Department until 1967. Since then it has been a part of the extended empire of the Department of Transportation. Cases can be made for “no change” (stay in Transportation), some form of independent agency (the Federal Emergency Management Agency?), a move to the Justice Department, or—as suggested here—a move to Defense. There are problems with each alternative. Suffice it at this point to argue that the military benefits that should flow from the Coast Guard’s being fully recognized, and treated bureaucratically, as one of the five armed services match trends in Coast Guard duties, facilitates the vital national-fleet concept/treaty/plan, and should promote vital commonalities in C4ISR, \emph{inter alia}. (An interesting case can be made for moving the service to the Justice Department, but—considered overall—Coast Guard activities fit better within the Defense Department, or DoD.) The challenge, of course, would be to identify ways to reap the military benefits of Defense family membership while retaining the distinctively law-enforcement and regulatory, civilian-lifeguard, and resource-protecting character and ethos of this uniquely dual-focused service.
CONTEXTS
The future force structures of the Navy and the Coast Guard, and the terms and conditions of their relationship, will be shaped vitally by the several contexts that give them meaning. Four contexts dominate: domestic (resource), foreign, military, and (nonnaval) maritime. The fourth, “(nonnaval) maritime,” refers to a broad understanding of the uses of the sea; its inclusion here is essential, because it accommodates the many activities and trends by which are created public demand for coast-guarding services. While the Navy and the Coast Guard are complementary and can be fashioned at the margins so as to take cost-effective account of each other’s assets, they are nonetheless driven by the beats of very different drummers. At least forty or so of the world’s “navies” are really greater or lesser coast guards, but that is not, and cannot be, the case for the global superpower. In fact, as our argument explains, the military context for the U.S. Navy is changing in ways that accelerate its distinctiveness from the Coast Guard. This is not a criticism; it is just the way things are in the military realm at the beginning of the new century.

We should note that our current strategic condition is not exactly the first period in modern history when the greatest navy of the day had to cope with the challenges of what became a long political peace (to borrow from Edward Gibbon) and an accelerating technological revolution. C. J. Bartlett, in his now-classic study of British sea power in the four decades from the fall of Napoleon to the Crimean War, writes:

Many were the calls upon the [British] government after 1815 for cuts in official spending, for tax concessions, and for money for other purposes. A weak ministry would neglect such demands at its peril. Naval supremacy could not be sacrificed, nor would the country have allowed such a thing, but many savings short of that point might be exacted.

Earlier he advises that

the formulation of naval policy is perhaps best viewed as a fluctuating, intermittent, triangular struggle, with three main considerations, sometimes in opposition, determining the decisions. These considerations may be described as domestic, foreign, and naval.12

The Domestic-Resource Context
In long periods of peace, especially when the character of war is being rapidly altered by new technologies, strategic rationality rarely governs defense policy. Among the reasons why this should be so is the uncomfortable fact that strategic need is inherently contestable—whereas, there is by definition a shortage of convincing reality-tests to sort good ideas from bad ideas. Why does the United
States, with more than a nine-trillion-dollar gross domestic product, spend approximately $290 billion annually in its Defense Department budget? Is roughly 3 percent of GDP the “correct” resource allocation? Historical scholarship and common sense tell us that the overall level of U.S. defense expenditure in peacetime is driven by public mood. For a while, at least, the United States can afford to spend on defense literally whatever it takes to see off the foe of the moment.

Americans today assuredly could (and almost certainly should) choose to spend close to $400 billion. But this economic feasibility could be translated into political demand (voiced by the people’s tribunes in Congress) only by a mood swing, triggered by a shift in threat perception. Although the armed forces, including the Coast Guard, enjoy high public approval ratings, the nation is not about to write them a very much larger check—not, at least, until the need appears to be demonstrated by real-world events. Careful defense analysis probably can show quite plausibly that the country can and should spend that extra hundred billion a year on national defense, but to advocate as much is to “whistle Dixie,” absent some indisputable revelation that the military machine is broken or inadequate. (We won the Cold War and the Gulf War, didn’t we?)

The politics behind U.S. defense expenditure since World War II have produced an irregular cycle, a “wave train,” of surging and then falling outlays as the tide of popular alarm has risen and receded.\textsuperscript{13} If we ignore the direct costs of the Vietnam War, we find that there have been two surges—one truly great, the other merely major—in the U.S. defense budget over the past half-century. The first, a tripling of the DoD budget fueled by the shock of Korea in 1950, allowed the country to arm properly (and then some) to wage the Cold War seriously. That was the “wave” that bought the “real Strategic Air Command,” not to mention a nuclear-armed navy and army. The second financially significant “wave” began to rise in the final year of Jimmy Carter’s presidency and climbed usefully through and beyond the mid-1980s (the Ronald Reagan years), as plans and contracts eventually led to actual expenditures.

Great new ventures in U.S. national defense, including candidate revolutions in military affairs and military technical revolutions, may be conceived in the lean years between “waves” (think of carrier warfare, amphibious warfare, long-range bombing, armored warfare);\textsuperscript{14} to be transformed, however, from ideas into capabilities they have to catch and ride a “wave.” The Strategic Defense

---

*With a 600-ship Navy, 40 or so Coast Guard cutters were virtually an after-thought. With a Navy of 116 or fewer surface combatants, and in a world plagued with regional instability and strife, however, our 40 major cutters along with several hundred coastal patrol boats take on new significance.*

\textsuperscript{13} Admiral James M. Loy, USCG, 1997

---
Initiative, for example, missed the Reagan budget wave. The United States will deploy a multilayered ballistic-missile defense architecture on a scale suitable to the threat only if the idea rides some future swell in the defense-budget wave train.

The budgetary context of the future of the Navy and the Coast Guard, and particularly the future of their relations, is therefore as follows. First, the defense budget currently is in the trough between waves; the level is rising, but not significantly. (In today’s dollars, the DoD budget has declined from $436 billion in fiscal 1985 to $291 billion in 2001; the Navy budget has dropped over the same period from $150 to ninety-plus billion.) Second, there will be a new wave, or surge, in the budget the next time Americans are persuaded that they face a clear and present danger. That new wave will very rapidly raise all boats. The lesson of the early 1950s and early 1980s is that Congress and the White House will want to fund virtually whatever is ready to be purchased (and, of course, much else besides). If ballistic-missile defense is arguably ready, or ready enough, it will be bought on an impressive scale.

If this historically and culturally based argument has merit, it means both good and bad news for the U.S. Navy. The bad news is that until there is a fairly dramatic, and self-evident, deterioration in the security environment, ninety to a hundred billion dollars (in fiscal 2001 currency) will be the politically authoritative and enduring support level. It does not much matter who is in the White House: no administration is going to effect a step-level jump in dollars for the Defense Department, let alone for the Navy in particular, unless something significantly and unexpectedly nasty occurs in the world. The good news, however—if one dares put it that way—is that the defense budgetary wave train is assuredly still working. It is a thoroughly safe bet that sometime in the next few decades the country will decide—probably rapidly and on the basis of scant strategic analysis—that it must have a much larger, at least more capable, navy. The political challenge is to be ready to mobilize when the moment arrives and meanwhile to make such investment as is possible in the $90–$100 billion fleet, so as to minimize regret later on.

After all, strategic history has a way of delivering crises and wars for which countries and armed services are less than optimally ready. Recall that the U.S. Navy was ready for war by late 1943, not 1941–42, and that Hitler’s Kriegsmarine had been hoping to be “game fit” by 1949, certainly no earlier than 1946. Bartlett captures the point in his treatment of the Royal Navy after the Napoleonic Wars:

Respect for this principle [to maintain naval supremacy] did not prevent ministers from living upon “capital” accumulated in the past, from economizing in stores and new ships in the interest of domestic politics and finance, or from putting the immediate calls upon the Navy before future security. . . . Altogether it is doubtful whether
any power with a parliamentary constitution, with Britain’s insular position and a
history of territorial inviolability for so many generations, could maintain in time of
peace a fleet prepared in every particular for a large-scale war. The test should not be
perfection but the adequacy of the fleet to deal with an unexpected emergency: not
whether sufficient money and effort had been expended but whether economy had
been pushed too far.15

The Foreign Context
The second context is, confusingly, both a cause and effect of the first. What
Americans allocate to national defense is a function not only of events in the
outside world, the course of international history, but also of the policy they
choose to follow. Foreign policy can age far more rapidly than ships and aircraft.
The new carriers, destroyers, submarines, aircraft, Marine weapon systems, or
maritime security cutters that we debate today may well be staple elements of
the national fleet beyond mid-century, meeting the strategic demands of foreign
policy (and domestic policy, for the Coast Guard). But what do we really know
about the “out-decades” of the century we have just entered? Perhaps more to
the point, does it really matter what we do and do not (and cannot) know? Can
we simply register a few grandly sweeping political assumptions about the U.S.
role in the world of tomorrow and then, having checked that box, sail smartly on
to what we find genuinely interesting—a possible revolution in naval affairs,
ship design, network-centricity, the perils of combat in the littoral, and so forth?

Because we are speaking of the future of the Navy and Coast Guard, we must
consider the foreign-policy driver of U.S. defense in a strictly functional way.
Just as naval analysis needs to be contextualized by the perspective of maritime
strategy (a healthy dose of Julian Corbett), so strategy requires a political con-
text.16 Are the Navy and Coast Guard to be so confident in political assumptions
about the future of world politics and the U.S. role (and derivative missions)
therein as, in effect, to regard foreign policy as a given? We have characterized
domestic-resource context as an irregular wave train, with occasional great and
not-so-great budgetary surges, triggered by the ringing of security alarms.
Can the Navy shape long-term acquisition programs (such as the DD 21) with
only generic foes in mind? Surely, distinctive geopolitics might imply distinctive
ger of war (for instance, China’s strategic geography is dramatically different
from Russia’s). If so, defense planners today should have no difficulty ac-
knowledging, at least in principle, that the relative importance of naval power in
joint warfare should vary with the identity, and hence geography, of the enemy;
and that the utility of different kinds of naval capabilities should vary with
the character of combat specific to adversaries and places. Geography does mat-
ter.17 It may have seemed to be overridden when U.S. naval power reached into
landlocked Afghanistan in 1998 to try to touch a rogue among rogues, but as a strategic enterprise the exercise was not entirely satisfactory. Technology is not synonymous with strategic effect.

Some of the contemporary speculation about the future of the U.S. Navy is highly intriguing but deeply flawed—specifically, by apparent unfamiliarity with the enduring lore of global strategy. Jan Breemer, a leader of the “Monterey school,” for example, argues that “naval strategy is dead,” while others note that ours is a “golden age of United States sea power.” On the first assertion, one might be moved to quote Mark Twain’s wry comment that reports of his death were greatly exaggerated; as for the second, all golden ages without exception fade and pass away. Characteristically, golden ages bear the seeds of their own demise. They glitter most brilliantly when they have already endured many years and seem likely to endure for many more—but are in fact nearing dissolution.

The Navy Strategic Planning Guidance records that “the Navy assumes that no peer competitor on a global scale will arise prior to 2020.” That assumption is probably correct, though the qualifying phrase “on a global scale” merits watching. Nonetheless, consideration of the foreign and the (nonnaval) maritime (see below) contexts of the twenty-first century reduces confidence in the “golden age” argument. There are at least two torpedoes in the water racing to destroy that view. First, the whole history and logic of international relations tells us that the United States will not be permitted to exercise unchallenged a benign hegemony on behalf of the existing global “order.” One may attempt to argue that the present course of events shows a profound transformation away from the dangerous competitive practices of the past, but the evidence is weak and assuredly inconclusive.

Second, and as a strategically logical corollary to the first point, it would be absurd to suggest that the greater among America’s enemies in this new century will concede, without contest, to U.S. sea power—writ large, the right to use the seas at will. It is about as prudent a thing to suggest today as it would have been in 1815.

Let us be crystal clear: If it continues to be important, indeed vital, that the United States be able to use the seas for commercial and military purposes, then its rivals and enemies must inevitably be motivated to contest that ability—directly or indirectly, symmetrically or asymmetrically. Little comfort should be drawn from the fact that the U.S. Navy is currently the “last man standing.” Resources, time, and strength of political motivation will find a way to challenge that status. Mahan was as correct for 1890 as he was for 1990, and prospectively for 2090, when he wrote, “Notwithstanding all the familiar and unfamiliar dangers of the sea, both travel and traffic by water have always been easier and
cheaper than by land.” Whatever technological erosion Mahan’s claim has suffered with the passage of 110 years has been only marginal.

Of course, the U.S. Navy is not opposed today by any other navy with global pretensions. But that is not a profound, or even particularly significant, observation. The most basic reason for this “golden age” of U.S. maritime command is eminently temporary: world politics has been shaking down from the Cold War, is in an interwar period of uncertain duration, and no polity or coalition now existing has both the political incentive and the readily mobilizable base to challenge the U.S. maritime imperium. The *Navy Strategic Planning Guidance* makes its claim only to the year 2020. It is highly improbable (though not inconceivable) either that the United States will be permitted to enjoy a Second American Century untroubled by serious competition or that a major foe, probably a coalition of foes, will be content to concede the high seas as a sanctuary and highway for such commercial and military activity as the United States chooses to conduct. The current understandable focus on operating “from the sea,” “forward from the sea,” and even “beyond the sea” should not lead to a self-indulgently static view of the future of international politics. American defense theorists have to avoid intellectual capture by the “grammar” of contemporary military affairs at the expense of longer-range strategic thought that is properly contextualized politically.

Let us close this part of the discussion with a historically based caveat for those susceptible to “golden age” fallacies. Such errors misled both the imperial Athenians and Britons, and they have induced others, like the continentalist Spartans and Germans, to acquire fleets (leading the latter to a strategically unsuccessful outcome). In this century, the greater among America’s foes are not likely to concede all operations in the entire maritime realm beyond coastal waters to the authority of the U.S. Navy. This brings us to the inherently contestable military context of guesswork about future warfare.

The Military Context

Navy–Coast Guard relations in the framework of the national-fleet concept have to be shaped, in their military aspects, principally by the military future of the Navy. That future has implications for the Coast Guard, but national defense is only one among the roles of that service. We have to be careful lest in a search for cost-effective complementarities between Navy and Coast Guard we underplay the vital point that the first duty of the latter is excellence as a coast guard. If the U.S. government decided, foolishly, that the overriding job for the Coast Guard was to assist the Navy, it would be sensible to fold the Coast Guard wholly into the larger service.
The U.S. Navy takes its tune, or at least should, from control (even command) of the high seas. Evolution in the grammar of sea war, and of war generally, will from time to time suggest shifts in emphasis among subsurface, surface, air, space, and indeed land-based naval assets. What is significant as a matter of broad guidance for planning, however, is that the first duty of a navy, in this case the premier navy, is to control sea lines of communication—to allow or deny access to the sea, thence across it, and finally to the land, where humankind lives. Whether or not today we judge it probable that any foe of the United States will try to build a high-seas fleet over the next half-century that might mount a challenge for maritime command, thereby hazarding a Trafalgar or Midway, is not at all the point. The true point is that the more powerful of the nation’s foes in the future are certain—yes, certain—to need ways to defeat or sideline the U.S. Navy. Much as the European Nato members were always unenthusiastic in the Cold War about strategies geared to “win” a war, as if on points, in Europe, so America’s enemies would prefer to do more than just deny to U.S. military power their own littorals. Sophisticated fortress design usually included outworks, to act as attriting and distracting breakwaters. Contemporary China, for example, like imperial Japan before it, understands the concept of area denial very well indeed.

The military context for the U.S. national fleet comprises so rich a set of topics that it can be treated here only in summary fashion. First of all, the great RMA debate of the 1990s is over. By historical analogy, it is in the condition reached in the nuclear debate by the early 1960s: the debaters are out of fresh inspiration; the intellectual offerings have become overelaborate; the real world of strategic history has provided little new stimulus; and antithesis has done its work, leading to some synthesis. Scholars differ on whether an information-led RMA is a seismic military revolution, meaning a wave to be ridden but not directed, or merely a military technical revolution, to be “made” by dedicated revolutionaries in the face of atavistic forces. We must not forget that these terms are no more than the conceptual inventions of imaginative people; we should not take them too seriously. Competent theorists always can invent and deploy exciting-sounding concepts to explain the necessity for the navy they prefer. The RMA family of ideas is inherently biased in favor of the novel. Also, it is worth noting that the recently advertised “revolution in naval affairs” has less to do with

Since the last Quadrennial Defense Review, I’ve said—and believed—that a force of 305 ships—fully manned, properly trained, and adequately resourced—would be sufficient for today’s requirements—within acceptable levels of risk. But . . . mounting evidence leads me to believe that 305 ships are not likely to be enough in the future.

Admiral Jay L. Johnson, USN, 1999
military science than with a political reality: that the (somewhat) blue-water Soviet naval challenge to U.S. maritime superiority was terminated by radical change in its domestic political context.

Second, a U.S. Navy planning “to boldly go” (to use a split infinitive institutionalized by a famous television series) where great navies have feared to sail, into coastal waters—at least to operate forward from, and beyond, the sea—needs to be alert to a logical paradox of conflict that Edward N. Luttwak has all but made his own. In the same way that a United States apparently omnicompetent in conventional warfare motivates enemies to find, if they can, asymmetrical strategic “equalizers” (such as weapons of mass destruction, terrorism, and political warfare), so a U.S. Navy shaped significantly for littoral warfare should expect foes to investigate other regions in which to seek its embarrassment.

Third, whether in coastal waters or on the high seas, the navy of a superpower that aspires to protect commerce and international order globally has no responsible choice other than to pursue excellence virtually wherever military science takes it, however serendipitously. (Remember the trireme; recall HMS Dreadnought and then Queen Elizabeth.) Athens and Britain had structurally similar problems in translating naval strength into strategic advantage, but they were both entirely sound in policies of maritime supremacy. Such supremacy alone did not ensure their security—which is why we read Corbett as well as Mahan—but it was the keystone of their strategies. The U.S. Navy in the twenty-first century, following Pericles and Jackie (properly, Admiral Sir John) Fisher, has to be and remain unarguably “second to none.” In an important sense, it does not fashion itself to defeat the navies of today’s (absent) enemies; rather does it plan to control the sea, and especially the high seas—failing which any discussion of littoral warfare is strictly moot—against all comers, in all combinations, and with any style of combat. Politically and strategically, this is the nature and character of the U.S. Navy with which the Coast Guard must cooperate in a national fleet.

Fourth, twenty-year expectations should not drive forty-year capabilities. The current naval debate is pervaded with some doubtful orthodoxies that are probably just plausible fallacies. For example: naval strategy is dead; naval warfare is now just warfare, not even an especially distinctive category of it, and accordingly naval vessels are simply vehicles more or less interchangeable and competitive with those based on land (or in space); and blue water is “American” by history’s definitive strategic verdict. Perhaps victory in three great wars (two hot, one cold) prompts this supposition. Unfortunately, the most recent victory—all but an accident in geopolitical and geostrategic terms (the demise of the erstwhile other superstate)—is being interpreted as a lasting condition in
its naval implications. Devotees of the proposition that the high seas are permanently American should ask themselves on what basis they judge themselves superior in predictive talent to the statesmen of the 1920s who failed to foresee the 1930s, or of the 1970s and early 1980s who missed the grand Soviet collapse of 1989–91.

Finally, technological change drives tactics, shapes operational art, and can imply strategy. However, certain inherent strengths and limitations of sea power, and naval power, have proved remarkably resilient in the face of technological, social, economic, and political change. The demise of navies has been much anticipated but appears no more imminent today than it was half a century ago. In fact, as Sam Tangredi has argued (and we will suggest in the maritime context), current trends in globalization are augmenting, not shrinking, the relative importance of sea power in general, and therefore of navies and coast guards in particular. Modern military science has not so homogenized military affairs that physical geography is strictly yesterday’s dimension of war and strategy. Ships can loiter menacingly as no other form of military power can. The (admittedly expensive) self-defending, mobile airfield that is a fleet carrier (and its escorts) means—to paraphrase the tag line of a once-famous 1970 movie—never having to say one is sorry to, ask political permission from, or consult the sensitivities of, a local friend.

Of course, there is much that the Navy and Marine Corps cannot do, or cannot do well—as both are continually reminded by the Army and Air Force as well. That, we must add, is one reason why there is a U.S. Coast Guard. We can illuminate with a four-way discriminator what each of the four environmentally specific services—and, distinctively, the Coast Guard—brings to the banquet in contrast to the Navy. We ask:

• What each can do uniquely
• What each can do well
• What each tends to do poorly
• What each cannot do at all.

As best one can tell—and the historical evidence is truly impressive—naval power is in no danger whatsoever of being overtaken by a technological, or any other, revolution. Mahan and Corbett are classics not only of sea power but also of strategy. Much of their explanation of the strategic utility of navies and of sea power is as valid today as it was a century ago.

The (Nonnaval) Maritime Context
The three contexts discussed thus far have leaned heavily toward consideration of the future of the Navy in the national fleet alliance. The fourth context shifts
the focus to the Coast Guard. It so happens that we can be as certain about trends in the nonnaval uses of the sea, and therefore in the demand for distinctively coast-guarding services, as we are necessarily uncertain about the major problems that could test the Navy in combat. The guesswork involved in naval planning is nearly absent from forecasts bearing upon the Coast Guard—because the primary focus of the service is not the national defense duties legally laid upon it but marine safety, maritime law enforcement, and marine environmental protection. Whatever (say) China attempts, the U.S. Coast Guard will have to try to cope with an expanding domain and the growing severity of challenges in its unique role as the country’s policeman at sea (and on shore). What do we know about the future of non-national-defense demands upon the Coast Guard?

The pressure on America’s maritime frontier is increasing and will continue to increase. It ranges from a huge foreseeable growth—probably a tripling—in the volume of legitimate maritime trade by 2020 (95 percent of U.S. exports and imports now move by sea); a boom in illegal migration by sea; an increase in marine drug smuggling; and greater demands on the resources of the sea, including fisheries and mineral deposits. The general picture of growth is clear enough. Those who guard the maritime approaches know that more and more is required of them, aside from their also-growing list of defense-related duties.

The bare statistics on Coast Guard activity are awe inspiring. The American public is reminded of the Coast Guard when a white hull with the orange and blue stripes is seen in the context of an air crash at sea or the interception of a boat with illegal drugs or immigrants. The big picture, however, is that the service responds each year to approximately fifty thousand distress calls, saves perhaps five thousand lives outright, and provides help to a hundred thousand other people in emergencies. Every year, roughly eight hundred Americans lose their lives in marine accidents. Seventy-eight million Americans a year now use recreational vessels, a figure predicted to rise by 65 percent over the next twenty years. Relatedly, cruise ship fleets will double in that period; individual vessels will carry five or six thousand people and displace up to 142,000 tons, fifty thousand more than the carrier Nimitz. Consider the sheer scale of another key current Coast Guard activity: each year the service conducts safety inspections of thirty-four thousand U.S. vessels, nineteen thousand foreign vessels, and seven thousand fixed marine facilities on and off shore.

The relevant trends are almost entirely in the direction of heavier burdens, and the reasons are all but self-evident. For one, the world’s population is likely to expand by a staggering two billion by 2020. For another, the United States is a giant economy (nine trillion-plus dollars) with gigantic maritime responsibilities, including guardianship of the world’s largest offshore exclusive economic
zone (or EEZ, at 3.36 million square miles); U.S. economic performance is critically dependent on the high seas and its own coastal and internal waterways. Everywhere we look we find an increase in pressure on marine resources and, of course, on those charged with their protection. For just one example, fisheries law-enforcement boardings by the Coast Guard increased from 9,440 in 1994 to 14,173 in 1998. No fewer than twenty-six thousand commercial fishing boats operate in U.S. waters. The context of pressure on the Coast Guard, in fact, most vividly emerges from the juncture of these two factors: the predicted growth in world population translates to an annual demand for perhaps 110 million tons of protein from fish, while only eighty-two million is likely to be prudently harvestable. It follows that the Coast Guard, charged with protecting the world’s largest EEZ, with 20 percent of the world’s estimated fish resources, is going to be kept very busy.

The litany goes on. The next twenty years will see an expansion of oil imports (by sea, of course) to the United States, from today’s 40 percent of domestic consumption to 66 percent of an even larger total demand. There will be more shipping, larger and faster ships with smaller crews, and many more people taking their leisure on, and traveling (by ferry) through, U.S. ports and waters. With regard to drug traffic, for example, at least 70 percent of the total drug flow into the United States now travels part of the way by sea.

The fact is that each of the three segments of the Coast Guard’s portfolio—justice, transportation, and defense—is registering noticeable growth. The service, uniquely, is responsible for law enforcement at sea (being the only armed service exempt from the restrictions of posse comitatus), leads in marine safety in all respects, and has a noteworthy clutch of maritime defense assignments. Therefore the significance to the nation of the Coast Guard is in effect that of the sea itself. Only those who do not appreciate how vital is the sea to the safety, prosperity, and well-being of the United States will find difficulty in understanding why an obsolescing Coast Guard matters deeply.

A NATIONAL FLEET
As noted at the beginning of this article, the Navy has entered the new century perilously thin in numbers of surface warships, a condition currently programmed to worsen markedly. Pessimistic projections show as few as ninety ships (all “high-end”) by 2020. The service has descended in a decade from a total of 592 ships to 316 today and is headed below three hundred. But times alter needs. The Navy has looked hard at what it may require to enforce access from the sea through the coastal waters of enemies, and it has unsurprisingly concluded that numbers matter.
Today the service is examining ways to square the circle—how it can increase the number of its surface combatants while carrying through a revolution in ship propulsion (integrated electric drive, which enthusiasts claim should prove as significant as the shift in the nineteenth century from sail to steam), all without breaking the bank. The much discussed, though as yet little endorsed, STREETFIGHTER concept of small but highly competent coastal warships is a notable sign of the times. By the Navy’s own sensible admission, the plainly predictable scarcity of surface warships and attack submarines threatens the integrity of policy and strategy. Mean, potent STREETFIGHTERS may or may not be part of the strategic solution, but the thirty-five to forty maritime security cutters that the Coast Guard plans to acquire over the next twenty years have to be.

The present high and medium-endurance cutters, patrol aircraft, and communications systems of the Coast Guard are generally old, becoming obsolescent, and in some regards already inadequate. In the apposite words of the Commandant of the Coast Guard, Admiral James M. Loy, “Our ships and aircraft that operate offshore are among the oldest of all the world’s fleets of similar platforms—38th of 41. The personnel and maintenance costs of keeping these ancient craft running grow increasingly prohibitive.” The point is not that obsolescent equipment cannot do as good a job as more modern equipment—that is obvious—but that it will increasingly be unable to do the job at all. Smuggling craft, for example, already outrun or otherwise evade Coast Guard vessels, and their radar cross sections defeat the old search radars on the cutters. Similarly, the maintenance problems of aging ships and aircraft affect endurance on station, especially far from shore.

The need to recapitalize the deep-water assets of the U.S. Coast Guard draws urgency, then, both from the ever-more-challenging scale and quality of traditional coast-guarding tasks and from the need to operate effectively with the other U.S. armed services. The sharpness of this problem has been highlighted usefully by the Coast Guard’s Assistant Commandant for Operations, Rear Admiral Ray Riutta. He confides that “the biggest challenge I have...is to make sure that I do not field assets that are so slow, defenseless, and technologically outdated as to be albatrosses around the necks of the Navy’s forces with which we sail.” To underfund (as today) the Coast Guard while its duties expand is to produce an entirely predictable result: its equipment and people are worked harder than is wise; retention, equipment life, and readiness rates suffer; and a crisis gathers.

To sustain operational tempo in the near term, spare parts are cannibalized from working equipment. That practice is now standard for the Coast Guard’s fleet of thirty C-130s. Some planes are kept flying, but at the cost of reducing an
already lean force structure, demoralizing overworked personnel, and lowering retention rates of experienced service members. (Between 1994 and 1999, the availability of Coast Guard C-130s dropped from 80 percent to an unacceptable 60 percent.) Cannibalization means that at least one of the aircraft assigned to each of the Coast Guard’s five air stations is likely to be a “hangar queen.” These planes, of course, are in no sense “extra”; stripping them for parts means patrols not conducted or backup not available; the consequence can be lives needlessly lost at sea and laws not properly enforced. The issue is easy to explain. Admiral Loy simply states the facts when he says: “The unfavorable trends in aircraft availability, parts inventories, and crew experience challenge our ability to provide mariners in distress with the rescue services Americans have come to expect.”

The national security community typically shows as little interest in the Coast Guard as did the Clinton administration (understandably, the jury is out on the Bush administration). Many of the problems described here lend themselves to elementary and inexpensive solution. Again to quote Admiral Loy, speaking on the same occasion, “Many of our readiness issues are the sort of problems that really can be solved by throwing money at them. Twelve or thirteen million dollars to restore our parts inventories to where they were a few years ago would be a nice place to start.” Both intellectually and financially the Coast Guard operates beneath the notice of most political leaders. In 1998 the White House approached Congress for an emergency supplemental budget augmentation to improve the readiness of the armed forces. The administration’s request did “not include a nickel for the Coast Guard,” even though the service had participated in its development and shared the readiness problems of all the other services.

It would be a serious mistake to assume that the crisis described may produce, at worst, a graceful degradation of performance—“merely” more lives lost at sea, more pollution not prevented or punished, more drugs brought in by “go-fast” boats, and more fish stocks illegally run down. A persistently underresourced, aging Coast Guard could also mean large and exceedingly unhappy events: another Exxon Valdez disaster, a super-cruise ship aided too late, or a nuclear device transported by sea undetected into the United States.

Aside from all this, the Coast Guard has to be ready to place itself in harm’s way in active war zones. It is not the mission of the service to win and keep control of the sea; that is the job of the Navy. However, the Coast Guard does have to be able to operate seamlessly with the Navy and the other armed services. As the Navy heads for mastery of information-led network-centric warfare, it must help the Coast Guard to function within that framework. The C4ISR regime needs to be common to all of America’s services, not just a “high-tech” four.
The emerging picture is one of national security under increasing stress at sea. The Defense Department and the Navy admit—in the national fleet “treaty” signed with the Coast Guard—that the fleet is too small and that accordingly the Coast Guard has to reequip itself for duties farther offshore. Forty-one deep-water cutters were a useful adjunct to the (almost) six-hundred-ship Navy of the 1980s. Today, as the Navy copes with austerity and technical revolution by concentrating its capabilities in fewer ships, optimized for high-intensity conflict, those cutters assume a new significance. By 2020, the Navy will have no frigates left, and all its destroyers will be large and expensive. There will be no small, general-purpose warships—a yawning void between the fourteen Cyclone-class (PC 1) patrol boats on the one hand and the fifty-seven Arleigh Burkes and the new DD 21s on the other. This trend could change in the future, particularly if the service pursues some variant of the STREETFIGHTER coastal warship, but certainly it is not going to change rapidly. The Navy intends to replace no fewer than eighty-eight Oliver Hazard Perry frigates and Spruance destroyers with thirty-two DD 21s (really mini–Arsenal Ships, close in concept to the monitors of yesteryear), beginning with the first DD 21 in 2008 and about three ships a year thereafter. By 2020, as noted already, Navy surface warships might well number scarcely more than ninety.

The interconnecting parts of the compounding maritime crisis, however, suggest a workable solution. For once, the law of unintended consequences should work in favor of U.S. national security. If the Navy expects to be short of, or even lack entirely, frigate-like warships of the future, the Coast Guard can help fill the gap; it need only reequip itself in a way that it ought to anyway for the better conduct of nondefense duties. The concept is that of a national fleet: the idea that the Navy and Coast Guard are complementary and synergistic, that they are especially cost-effective when yoked in tandem.

The concept is bureaucratically new, but it is well established in historical practice; no break with law or tradition is contemplated. The Coast Guard is legally an armed service, and its cutters have always been more or less well-armed, shallow-draft warships. The question is how much “war” equipment Coast Guard warships should be capable of carrying. That question has yet to be answered in detail, but as a matter of guiding principle it is answerable today; indeed, the agreed concept of the national fleet settles the most pertinent issue. In political fact, of course, matters are not so simple. At present the Coast Guard must compete within the Department of Transportation for respect and dollars; within the administration for provision within the balanced budget; and in the face of a Navy all too willing to offer cast-off equipment (frequently overage and therefore expensive to maintain, crew, and operate; or otherwise excellent craft like the Cyclones, inappropriate for the Coast Guard’s blue-water tasks).
The zone of reasonable disagreement about the Coast Guard’s duties and equipment is not extensive. There is no uniquely right number of Coast Guard cutters, no unquestionably correct timetable for their replacement, and certainly no unarguably optimal equipment for them. Probably it is intelligent to think about a flexible, mixed fleet of deep-water cutters, perhaps with “plug-in” capabilities. Not all cutters would be available for combat zones, though obviously the more that are equipped to survive in a hostile environment—at least under the wings of other armed forces—the better. Also, the sea services need not decide soon and definitively precisely how “warship-like” new cutters should be. What matters is that cutters and aircraft be so designed that modular adjustments of combat power can be effected as required over operating lifetimes that could extend into the second half of the century, and that Coast Guard C4ISR allow interoperability with all the other services.  

All of this brings us to the only point in the argument not extensively supported thus far—that the Coast Guard, as an armed service, should seriously consider a move to the Department of Defense. The case can be put negatively or positively, and it is quite persuasive either way. The negative argument emphasizes the increasingly poor fit of the Coast Guard in the Department of Transportation. Too many of its missions and tasks—aside from safety of maritime passage, a core concern of the Coast Guard—lie far outside the charter, interests, and even culture, of Transportation. This is not a criticism of that department but an acknowledgment that the Coast Guard and its current administrative parent have evolved in different directions.

The positive argument for relocation to Defense is that America’s security environment increasingly requires the four armed services in that department to operate against irregular and unconventional foes. At the same time, the varied missions of the Coast Guard are requiring it to equip for, and perform, operations more military in nature. Whatever may have been the case in the past, the total job of the Coast Guard today is more complementary to those of the other armed services than it is to the Department of Transportation.

None of this denies the uniqueness of the Coast Guard among the armed services. As noted earlier, it is exempted from posse comitatus, and it is the federal policing agency at sea and on the coasts. Coast Guard people are lightly armed military professionals, many of whose tasks are notably nonmilitary (ship inspection, port safety, marine resource protection, safety of navigation, and so forth). However, an unmistakable practical convergence is under way: the Coast Guard’s defense mission is growing, while the national security agenda of interest to the Defense Department is widening.
Of course, there would be problems for the Coast Guard within Defense. However, the difficulties look distinctly manageable; they should be dwarfed by the advantages. The Coast Guard probably cannot fit snugly into the Defense Department (or anywhere else), but it can fit well enough there, provided the problems are understood and addressed. For instance, the essential recapitalization of the Coast Guard must meet distinctive Coast Guard operational needs; it must also capture complementarities and synergisms with the Navy. That would surely be easier to arrange within Defense than from outside. The material and strategic advantages, in general, could be huge. In Defense, the Coast Guard would be perceived and treated as what it truly is, the country’s fifth armed service.

Further, it would be able to participate more effectively in logistical, procurement, and personnel arrangements. Recent, indeed current, experience makes plain that the skilled people who are the Coast Guard increasingly suffer adverse disparities with respect to Defense Department personnel in benefits and allowances (though not pay). This fifth armed service has to try to recruit and retain professionals in competition with other services better able to look after their people; as it is, the Coast Guard can make up the difference only by siphoning money out of the equipment and operating budgets. In the Defense Department, the Coast Guard’s people would be treated as what they are, members of the U.S. armed forces.

The risks in all this are not trivial, but neither are they truly daunting. Above all else, a Coast Guard within DoD would need to preserve its distinctive warrior-policeman culture. There would be some areas of disagreement with other services. For example, it is probable that the Coast Guard’s expanding role as guardian of marine resources would fuel tension with the Navy, whose priorities have been rather different. So be it: America needs both perspectives. There is every reason to be optimistic that the service could retain its (somewhat) policeman-like reputation and yet function, and be seen to function, as a full member of the armed forces team. Perhaps a civilian assistant secretary of defense for the Coast Guard would be useful in achieving that balance.

Alternatives such as relocation to the Justice Department or conversion to something like the Federal Emergency Management Agency are interesting, and the arguments for them have merit, but they do not yet compete seriously with the Defense option. Looking to the decades ahead, the Coast Guard inevitably will be more of an armed service with a difference than a maritime police force with a difference. At the least, it is time for this subject to be aired publicly.

The dominant conclusion is that the Navy and the Coast Guard are natural complements and that this complementarity should be expressed practically in a “national fleet” that is not mere rhetoric. It is surely beyond argument that the
Navy cannot often afford to devote exceedingly scarce (and frighteningly expensive) assets to operations other than war. For its part, the Coast Guard, limping along on between four and five billion dollars per annum, could not take up naval—which is to say war-fighting—duties even if it so wished (which it does not). Even the next great surge in the defense budget (ideally by that time including the Coast Guard), which would presumably put the Navy back in the frigate business, would not promote serious difficulty vis-à-vis the Coast Guard. A rising budgetary tide would raise all boats—gray hulls and orange stripes. If—realistically, when—the American public becomes alarmed once again about defense, both the Navy and the Coast Guard (for its national-defense duties) assuredly will be bolstered.

Those who have the leisure to read history and strategic theory, and who are not immersed in the short-term issues and crises of government, have a duty to try to sort the enduring from the ephemeral. Officials lack the time (and, having chosen to be officials, probably the inclination) for such reflection. The whole subject of Navy–Coast Guard relations is in need of just such conceptual, strategic, and historical contextualization. In the heat, excitement, and urgen-cies of interservice issues, with their large budgetary ramifications, significant truths can slip out of sight. One is that, like the Athenian and British navies of the past, the U.S. Navy has an international “ordering” function that obliges it to dominate the most demanding possible conflict. The Navy is therefore prudent in investing in naval supremacy on the high seas, even though no plausible competitor is in view today.

Some concepts du jour hide their real obsolescence behind fashionable buzz-words, novel-seeming theories of future warfare, and mood swings, either optimist-mic or pessimistic. Donald and Frederick W. Kagan deftly skewer one such flawed view:

The absence of “global peer competitors” does not make the world safe for the foreseeable future—it only makes it uncertain and difficult to understand. When Japan attacked Pearl Harbor in 1941 it was not a “global peer competitor.” Nor was Germany in either world war. Still less was North Korea in 1950, North Vietnam in the 1960’s, Saddam Hussein in 1990, or Slobodan Milosevic in 1999 when they unleashed attacks that compelled American military intervention. None of these states posed an imminent threat to the continental United States. The concept of “global peer competitors” is the thinking of the Cold War and is not relevant to debates about future American national strategy.47

When the nation considers carefully the domestic, foreign, and military contexts of Navy–Coast Guard relations, much that is uncertain can be settled. The proposition that the sea services of the United States should behave as

https://digital-commons.usnwc.edu/nwc-review/vol54/iss3/30
complementary parts of a national fleet is true to their several natures and functions. It need not provoke controversy.

NOTES


2. Some of the reasons why this should be the case are explained in Colin S. Gray, Modern Strategy (Oxford: Oxford Univ. Press, 1999), pp. 217–27.


5. Ibid., p. 3.


9. With regard to active duty personnel, though of course not equipment, the U.S. Coast Guard, at 36,230, is not far short of the roster of the Royal Navy, which (without the 6,740 marines) numbers 37,030.


11. On IDS, see Stubbs and Truver, America’s Coast Guard, esp. chaps. 4–5.


20. See Colin S. Gray, “Clausewitz Rules, OK? The Future Is the Past—with GPS,” *Review of International Studies* (UK), Special Issue (December 1999), pp. 161–82, which tackles the old question of whether there is, or can be, progress in international relations.


29. For an interesting theory on the measure of U.S. responsibility for the Soviet defeat (did the USSR collapse, or was it pushed?), see Norman Friedman, *The Fifty-Year War: Conflict and Strategy in the Cold War* (Annapolis, Md.: Naval Institute Press, 2000).

30. Tangredi, “Beyond the Sea.”


33. On STREETFIGHTER in strategic context, see Arthur K. Cebrowski and Wayne P. Hughes, Jr., “Rebalancing the Fleet,” *U.S. Naval Institute Proceedings*, December 1999, pp. 31–4; and Hughes, “22 Questions for Streetfighter.” The eponymous “streets” are the enemy’s home waters.


35. By “deepwater” the Coast Guard refers to tasks conducted more than fifty miles from shore.


38. For example, when dealing with an article submission of mine on the Coast Guard, the editorial staff at the *National Interest* initially did not know even truly basic facts about the service. The journal had never published an article about the Coast Guard and probably had never previously even considered one. I was given the impression that the journal viewed the subject as so extraordinary as to be near the exotic! See Colin S. Gray, “Keeping the Coast Guard Afloat,” *National Interest*, Summer 2000, pp. 83–7. In some ways this episode speaks well for the journal, since it proved willing to step “outside the box” of what it had regarded typically as significant policy topics.

40. “Go-fast” boats account for 85 percent of maritime drug smuggling into the United States. Their use increased tenfold from 1995 to 1998. See John G. Roos, “Drug Busters,” Armed Forces Journal International, October 1999, pp. 34–40; Mike Emerson, “Coast Guard Helos: A Call to Arms,” U.S. Naval Institute Proceedings, October 1999, pp. 30–3; and Bruce Stubbs, “Launching the Real Maritime Security Force,” U.S. Naval Institute Proceedings, October 2000, pp. 72–5. “Go-fast” boats are usually thirty to forty feet long, have a couple of powerful outboard motors that can propel the vessel at speeds in excess of forty knots, and are austere in the extreme. They are designed, or adapted, simply to carry perhaps a metric ton of drugs at very high speed on runs of several hundred miles. Because these boats can outrun Coast Guard cutters, and because the Navy lacks legal powers of arrest, by far the best hope of stopping them is by nonlethal interdiction from the air.


42. The Navy’s proposal to increase the fleet from its current level of 316 ships to 360 was supported publicly by then–Secretary of Defense William Cohen. See Tangredi, “Beyond the Sea,” p. 21.


44. Lest anyone be confused about this matter, the relevant legal authority for the Coast Guard is as follows: 14 U.S. Code 1 specifies the Coast Guard as a military service; 14 USC 2 requires the Coast Guard to be ready to operate (and 14 USC 3 requires the Coast Guard so to operate) in the Navy in time of war; 14 USC 141 authorizes the Coast Guard to assist the Department of Defense in performance of any activity for which the service is specially qualified.

45. Kim Malmin et al., Future Coast Guard Cutter Study: Candidate Cutters and Their Costs, CRM 96-91.100 (Alexandria, Va.: Center for Naval Analyses, August 1996), is fundamental.

46. For an example of such divergence, see Rachel Canty [Lt., USCG], “The Coast Guard and Environmental Protection: Recent Changes and Potential Impacts,” Naval War College Review, Autumn 1999, pp. 77–90.

Dallas Isom’s article “The Battle of Midway: Why the Japanese Lost” [Naval War College Review, Summer 2000, pp. 60–100] is laudable for its use of Japanese sources and for the interesting points it raises. In particular, we applaud Isom’s interviews with Japanese survivors, which contribute new and useful information regarding Japanese aircraft rearmament procedures. This new data is crucial to building an accurate account of the events that transpired aboard the Japanese carriers on the morning of 4 June 1942. However, in our opinion, Professor Isom’s arguments appear to rely too much on a rather rigid (and highly debatable) interpretation of Japanese communications: namely, exactly when did Vice Admiral Chuichi Nagumo receive transmissions from scout plane 4, launched by the cruiser Tone. In addition, while Isom’s rearmament information (which he puts forward as central to Nagumo’s inability to launch his anticarrier strike before being fatally attacked) is clearly important to understanding the Japanese side of the battle, we feel that he did not carry his operational analysis far enough. As a result, we cannot accept his conclusions.

At the time of Isom’s writing, we were engaged in a reappraisal and rewriting of the Japanese account of Midway as a result of our own work in identifying wreckage from the carrier Kaga. A key part of our approach was to build an accurate model of the operations of the Japanese carrier striking force. As we will show, the disaster that befell the Japanese carrier force hinged neither on whether Nagumo received...
Tone 4’s message at 0740 or at 0800, nor on how quickly the armorers in the Japanese hangars could do their work. Rather, the fate of the Japanese Mobile Striking Force (Kido Butai) revolved around what was happening on its flight decks. Armed with a knowledge of Japanese carrier doctrine, as well as the operational information from the battle, one can reach an accurate assessment of the state of readiness of Nagumo’s force at the time of the climactic American attack without resorting to Isom’s indirect method.

A major error in the Isom article is that it repeatedly misstates what aircraft actually were on the Japanese flight decks. In several places, Isom presumes that the Japanese antiship strike force was on the flight decks when it is demonstrable that those aircraft were still in their hangars. This presumption carries forward the (mistaken) conventional view that the Japanese had spotted their antiship strike force on the flight decks shortly after the initial Midway attack force had taken off and had only briefly moved these antiship strike aircraft below for the purposes of rearming them or recovering the first-strike wave. As will be shown, this is the root of the fundamental misunderstanding of circumstances at the time the carriers were struck.

In fact, both Japanese doctrine and the operations of the Japanese combat air patrol (CAP) fighters would have kept the reserve strike planes securely below in their hangars until they were definitely needed. Not only that, but because of Japanese hangar design, the window of time necessary to lift, spot, and launch the aircraft was substantially longer than has been previously understood. As a result, given the frenetic nature of Japanese CAP operations after about 0800 (particularly aboard Akagi and Kaga), it is unlikely that many of these second-wave aircraft were ever spotted on the flight decks before the fatal American dive-bomber attack commenced at 1020.

This point cannot be overemphasized, because from the conventional belief of what was on the flight decks flow nearly all Western interpretations of the battle. To put the matter succinctly, at the time Akagi, Kaga, and Soryu were struck, their flight decks were more empty than occupied. This is almost the reverse of the standard view, which has the Japanese flight decks packed with strike aircraft awaiting takeoff. There were aircraft on deck, but most were CAP fighters, not attack aircraft. There were also fewer aircraft on deck than is generally supposed. Though potentially startling, this is less a “revision” than a correction and careful restating of the existing historical record. We will show that official Japanese sources on the battle have been aware of this for some thirty years.
These errors in both the conventional Western interpretation and Isom’s article cannot be addressed without first developing a sense of how the Japanese conducted carrier operations. Unfortunately, standard English-language histories of the battle of Midway have not well understood Japanese carrier operations. From the common misperception that Japanese naval aviation was derivative of Western (primarily British) practice, Western writers typically believe that the Japanese carriers of World War II behaved much like their Western counterparts. In fact, they did not. Japanese carrier operations contained elements of both U.S. Navy and Royal Navy practices. However, as a result of differences in physical design and operational doctrine, by the late 1930s Japanese carriers fought in a fashion all their own. Without understanding these points of divergence, understanding Nagumo’s actions is likewise impossible.

Before the days of modern angled flight decks, a carrier flight deck could be doing only one of three things: spotting aircraft, launching aircraft, or recovering aircraft. To this list most American writers would be quick to add “parking aircraft” and “servicing aircraft.” However, it is important to understand that the Japanese avoided performing these activities on the flight deck. Japanese carriers did not use permanent deck parks in the fashion of the U.S. Navy. While temporary deck parks were established forward of crash barriers during recovery operations, they acted only as transitional “overflow” mechanisms until returning aircraft could be moved below decks. Furthermore, while the Japanese could service aircraft on the flight deck (Japanese carriers were equipped with refueling points around the deck edge, for instance), most fueling occurred in the hangars. Likewise, except for arming dive-bomber aircraft, the Japanese serviced planes in the hangar as well.

Japanese carrier design is also notable for its use of enclosed hangar decks. In contrast to U.S. carriers, whose hangars could be opened to the elements by rolling up metal screens along their sides, Japanese hangars were fully enclosed by storerooms, workshops, and crew spaces, with no natural ventilation. This meant that aircraft engines were never warmed up below. This is in direct contrast to American practice, where aircraft commonly were warmed up in the hangars, brought to deck, and immediately launched. These factors had important ramifications during the battle of Midway and imposed severe restrictions on Japanese operational tempos.

In order to spot a strike force on the flight deck and launch it against the Americans (assuming it was already armed), Nagumo’s ships had to perform a
complex series of operations. Some tasks could be done in parallel, some only sequentially, and each entailed fixed or variable time costs. These included:

- Bringing the aircraft up to the flight deck (sequential: approximately one minute per plane).\(^5\)
- Spotting the aircraft, unfolding its wings, and chocking wheels (conducted mostly in parallel with elevator movements, but five sequential minutes are added to account for these movements).\(^6\)
- Warming up engines (sequential: fifteen minutes minimum for the entire force).\(^7\)
- In the case of dive-bombers, arming the aircraft (largely concurrent with engine warmup, but could take as long as twenty minutes).\(^8\)
- Delivering final briefings to the pilots (again, mostly in parallel with elevator movements, five sequential minutes minimum for the entire force).\(^9\)
- Moving crew to planes and performing final preflight checks (five minutes).\(^10\)
- Launching the strike (sequential: fifteen to thirty seconds per plane).\(^11\)

Taken together, it is apparent that spotting a twenty-one-plane strike for launch would take around forty minutes total, and another five to ten minutes would be required for the launch.\(^12\) If the deck spot contained dive-bombers, the spotting time would be perhaps five to ten minutes longer, because these planes had to be armed during engine warm up. This timing is directly confirmed in official Japanese sources.\(^13\) The need to warm up engines on the flight deck, dictated by Japanese hangar design, reveals itself as a major hindrance to Japanese operational tempos. Unfortunately, warm-up could not be shortened—aircraft casualties were the inevitable outcome of slighting this activity, and needless losses had to be avoided at all costs.\(^14\) Thus, if Nagumo was to attack the American strike force, he needed to find an unbroken forty-five-minute window of opportunity on all four flight decks during which to spot and then launch his strike.\(^15\)

The final piece of the puzzle is found in the activities of the Japanese combat air patrol that morning. An examination of Akagi’s flight operations reveals the basic point:

- 0430—launch Midway attack force
- 0445—launch initial combat air patrol (three fighters)
- 0543—launch CAP (three fighters)
0655—launch CAP (three fighters)
0659—recover CAP (three fighters)
0710—launch CAP (five fighters)
0720—recover CAP (one fighter)
0726—recover CAP (one fighter)
0736—recover CAP (three fighters)
0750—recover CAP (two fighters)
0808—launch CAP (three fighters)
0832—launch CAP (four fighters)
0837–0900—recover Midway attack force plus three CAP fighters
0910—recover CAP (one fighter)
0932—launch CAP (five fighters)
0945—launch CAP (three fighters)
0951—recover CAP (two fighters)
1010—recover CAP (three fighters).

The other carriers were involved in similar activities, albeit at different times. The important point is that recovery operations absolutely required a clear flight deck aft. Aircraft could not be spotted aft while aircraft were landing, nor would they usually be spotted aft during fighter takeoff operations.

This information enables us to appreciate several things. First, it is clear that Akagi was very busy on the morning of the fourth. The constant American air attacks from 0700 onward necessitated a continuous cycling of the Japanese CAP fighters above Nagumo’s formation. It was very difficult to find a spot to squeeze in forty minutes of uninterrupted deck time to spot a strike, let alone launch it.

This is a crucial realization in determining what was possible aboard the Japanese carriers and in analyzing Nagumo’s decision-making process.

Second, upon closer examination it is apparent that Akagi’s 1010 CAP recovery dictates that there cannot have been many strike planes on its deck when it was fatally bombed at 1025. Fifteen minutes would not have been sufficient time to bring its twenty-plane strike to the flight deck, let alone spot them, brief the pilots, and warm up engines. Thus, the common belief that the American dive-bomber attack found the Japanese flight decks practically chock-a-block with strike planes revved up and waiting to take off is clearly untrue. The aircraft on deck were primarily CAP fighters.
As it turns out, this latter observation, while perhaps shocking to an American audience steeped in the popular lore of this battle, is directly supported by official Japanese sources. In Japan, a clearer picture of Japanese Midway operations began emerging in the 1970s, with the publication of the official Japanese war histories (known as Senshi Sosho) and other substantive works. Unfortunately, it is only very recently that some of these works have been translated into English. Senshi Sosho explicitly states that at the time of the attack, every Japanese carrier had its attack aircraft in the hangars; the only aircraft on deck were either CAP fighters or, in the case of Soryu, strike force escort fighters that were being launched to augment the CAP.

Third, this operational information casts doubt on whether Nagumo’s reserve strike force was ever on deck in the first place. The conventional wisdom has always been that when Lieutenant Joichi Tomonaga’s Midway attack force was launched, the reserve antiship strike force was immediately brought up to the flight decks and spotted. In fact, this would be contrary to typical Japanese operational patterns, which would have preferred to keep the strike in the hangars until needed. It is absolutely certain that this force was not on deck during the 0800–0820 time frame, when Nagumo was making some of his most crucial decisions. Attacking B-17s photographed the Japanese formation during this period, and their pictures of Soryu, Hiryu, and Akagi show no strike planes on deck, only a handful of fighters. This is corroborated by Japanese records that show the force launching additional CAP fighters during this time.

With this information in hand, we now turn to several of the assertions in Isom’s article. For instance, regarding Nagumo’s supposedly tardy receipt of Tone 4’s message and its dire implications for subsequent Japanese preparations, Isom states (p. 75), “It should thus be apparent that if the rearming operation was reversed at this point—at 0745—it would not have taken much time to restore the torpedoes on half the planes from which they had been removed and respot all the planes on the flight decks of the two carriers, perhaps only about thirty minutes.” In light of the need for forty minutes just to respot the strike, not to mention the time needed to rearm, this gives a grossly optimistic impression of Nagumo’s chances of launching a strike before Tomonaga’s returning force would begin to occupy the flight decks at 0837.

Isom later states (p. 77), “Thus, at about 0920 operations to respot the second-wave strike force on the flight decks could have begun, had the torpedo planes been rearmed with torpedoes. Had the countermand order been given at
0745, as the standard scenario holds, the torpedoes almost certainly would have been restored by 0920." The ability of the Japanese carriers to begin a respot at 0920, had they been rearmed or not, is highly questionable given the high tempo of CAP operations and their monopolization of the flight decks. At this point in the battle, Japanese records clearly show, the Japanese were alert to a constant stream of incoming American strikes. Until the coast was reasonably clear and his CAP well stocked with fresh fighters to last through the spotting process, Nagumo cannot have been expected to spot his strike. Akagi’s air officer (hikocho) in charge of flight operations, Commander Shogo Masuda, as well as the other experienced air officers around Nagumo (such as Mitsuo Fuchida and Minoru Genda) could not have helped impressing this point upon the admiral.

Even more questionable is Isom’s subsequent statement (p. 78) regarding the state of Carrier Division 1 at 1000. When the attack was over at 1000, “about ten minutes of work still remained to rearm the last division of Akagi’s torpedo planes, and even more time was needed for Kaga’s. The torpedo planes that had been rearmed were brought up to the flight decks, beginning around 0920, but at least a third remained in the hangar decks at 1000. By 1015, the rearming had probably been completed on Akagi, and the last torpedo planes were being brought up and spotted on its flight deck. Had the whole strike force been ready to go at 1000, it, along with Zero escorts, could have been launched during this fifteen-minute window between attacks on the Mobile Force.”

This is wrong on several counts. First, we know that Akagi landed a CAP fighter at 0910 and two more at 0951, meaning that even if there had been strike aircraft on deck at 0920 (which we think highly unlikely in any case), they had to have been moved back down into the hangar by 0951. Also, Isom clearly does not factor in the immutable time costs associated with spotting and engine warm-up—a “fifteen-minute window between attacks” simply does not suffice. In fact, in this case Isom also ignores Nagumo’s own estimate that the strike force would be ready at 1100, although a 1030 takeoff was hoped for, if things went well. Launching at 1000, though, for all the reasons cited above, was never even remotely in the cards, and Nagumo knew it.

Furthermore, the assertion that two-thirds of Akagi’s torpedo planes were on deck at 1000 is clearly wrong. Akagi had landed CAP nine minutes earlier, at 0951, and would do so again at 1010. Isom’s assertion is also directly contradicted by Senshi Sosho, which states that at the time of the 1025 attack all of Akagi’s
strike aircraft were in the hangar. This is further corroborated by Richard Best, lead dive-bomber pilot against Akagi, who states that when he dove for his attack, the only aircraft on deck were Zeros.22

Isom makes a different error regarding rearmament activities aboard Carrier Division 2 (Hiru and Soryu). He writes (p. 79), “[the strike aircraft] could be re-armed on the flight deck as well as in the hangar deck. [It appears that only half of each squadron was lowered to the hangar deck after the 0715 rearming order, thus saving elevator time.]” Furthermore, he states (p. 80), “Most, if not all, of those [strike aircraft] had probably been changed back to armor-piercing bombs by very soon after 0830; at least half of each squadron on Hiryu and Soryu was already on the flight decks at 0830.”

Again, his statement is at odds with the photographic evidence obtained between 0800 and 0815, which shows no strike planes whatsoever on either carrier’s flight deck. In addition, it is known that Soryu launched CAP at 0710, and recovered CAP at 0730. Hiryu recovered CAP at 0700 and 0740, launched CAP at 0825, and was recovering CAP again at 0840. Thus even if strike planes had begun to be promptly brought up on Hiryu after the American B-17s departed at 0815, they would have had to be stowed below again by 0840—the window of opportunity was not long enough to have performed warm-up and launch.

Isom repeats this error later when he states (p. 81), “At 0830, when Nagumo had to make a decision whether to launch an attack on the American force or postpone it, we have seen that he had ready no torpedo planes and no Zeros for escort. But he did have dive-bombers on Hiryu and Soryu available. They could have been launched fairly quickly.” Again, this is incorrect, as the B-17 photographs and other evidence incontrovertibly demonstrate. The Japanese dive-bombers were all in their hangars at the time and would have taken another forty minutes to put into action, even if they had been rearmed.

Isom’s concluding point, and his central thesis (p. 89), “considering how close [Nagumo] came to launching his attack before being bombed at 1025, every minute saved could have made a significant difference in the outcome of the battle,” is shown to be incorrect by the cumulative weight of the evidence at hand. It was not the inefficiency of Japanese communications (which is debatable) that doomed Japanese hopes in the battle but the inefficiency of Japanese flight operations. Nagumo was nowhere near ready to launch by 1025; in fact, he had probably barely begun preparations to do so. Even assuming Akagi had begun lifting its strike aircraft to the flight deck immediately after its CAP was recovered at 1010, Akagi would not have been ready to launch its strike for forty minutes more (1050) and could not have gotten it completely airborne before 1100. Kaga was in a slightly better state, having last launched six CAP fighters
at 1000, but it was hampered because of its larger torpedo bomber wing (twenty-seven aircraft) to lift and spot.

It is no coincidence that after the devastating American attack on the other three Japanese carriers, Hiryu’s actual operational tempo corresponds very closely with the hypothetical earlier timetable for Carrier Division 1 we have just put forth. The flagship of the aggressive Rear Admiral Tamon Yamaguchi can be presumed to have launched its own strike as quickly as possible after the 1022–1027 debacle. Yet in the event it only managed to get that strike aloft by 1057, directly supporting the idea that its own deck-spotting activities had barely begun when the other three carriers were being attacked. Indeed, if the conventional view were true, planes should have been quite literally zooming off Hiryu’s flight deck at 1025 even as its compatriots were struck. Clearly, they were not, and this point is often overlooked.

The picture that emerges from this analysis is of a rather conservative admiral operating within the constraints of 1942 Japanese carrier doctrine. In contrast to the standard American accounts that have the Japanese observing the feeble American attacks (poorly executed and delivered piecemeal) with a measure of contempt, the operational tempo of Nagumo’s CAP reveals something rather different. Regardless of whether they accorded the American attackers much respect in terms of technique, the Japanese command was certainly concerned about the aerial assault, both for its disruption of the force’s timetable and for the overt danger posed by the attacks themselves. Upon close reading of the Nagumo report and other Japanese sources, there is no question that Nagumo and his staff knew the peril they were in. One good hit on any of the carriers could have disastrous consequences, and each of the carriers had survived near-misses during the morning.

Considering this, Nagumo probably thought he was playing it safe—keeping the strike aircraft in the hangars until the worst of the danger was past, keeping the flight decks clear to support constant CAP operations, and repelling American attacks with the best weapon available, his fighters. When the attacks abated, he expected to spot a coordinated combined-arms strike force on the decks of his two carrier divisions and then deliver a crushing blow to his opponent. Unfortunately for the Japanese, their desire to launch an integrated
attack force from all four flight decks deprived them of flexibility in the face of the enemy.

Because of the remarkably small cannon-magazine capacities of the Zero fighter, defensive CAP operations necessitated frequent landing, rearming, and launching of engaged fighters. Nagumo clearly appreciated the danger in which he would place his ships during deck spotting of strike force aircraft, in that it created a window of time during which no additional CAP could be cycled. As a consequence, his options were more constrained than has been previously understood—spotting an offensive strike meant hanging his CAP out to dry for nearly an hour at a time when American attacks were constant. As it was, Nagumo’s defensive approach very nearly paid off, as only the final American attack delivered telling damage. Indeed, it can be argued that had Nagumo played it a little safer, by putting additional CAP aloft, he might have saved Akagi, Kaga, or Soryu from the American dive-bomber attack.²³

Japanese carrier doctrine of the time did not specify what to do when suddenly faced with an enemy force within the enemy’s striking range while one’s own armed and fueled aircraft were still in their hangars.²⁴ This was a doctrinal failing—although in fairness, both the Japanese and Americans were grappling with this issue. Later in the war, it would have been considered imperative upon detection of an enemy force to immediately launch as many aircraft as possible (whatever their armament) against it. If nothing else, this would get the aircraft out of the hangars, where they presented a dire threat to the carrier itself. Indeed, by 1944 the Japanese Combined Fleet developed just such a command for the signal books in the event of a sudden enemy detection.²⁵

In conclusion, we applaud Professor Isom for his efforts in bringing Japanese sources to the fore of the Midway discussion, as well as his presentation of valuable information regarding Japanese rearming procedures. It is also important to bear in mind that he was laboring under the conventional belief that the second-wave strike was spotted and ready to launch on the flight decks, rather than below in the hangars. However, his interpretation of Japanese operations focuses almost exclusively on what it took to arm an airplane and fails to account for the fundamentals of how the Japanese got that plane spotted on the flight deck and then into the air. The article also overlooks the relationship between defensive CAP activities and the inability to mount offense strikes. Its view of carrier operations is therefore both limited and at odds with a great deal of what we know to be true about how Japanese carriers actually fought. Without an adequate appreciation of these factors, a proper assessment of Nagumo’s command options and performance cannot be constructed.

Editor’s note: Professor Isom responds on pp. 158–63 of this issue.
NOTES

1. The authors assisted Nauticos Corporation in identifying a large section of wreckage from the *Kaga*, discovered at a depth of seventeen thousand feet in September 1999. As a result of this project, the authors are currently working on a forthcoming book that will examine in detail the operations of the four Japanese carriers at Midway, bringing new Japanese sources to light in the process.


3. It should be noted that our comments pertain to early-war Japanese carrier design and doctrine. As such, some of our remarks may, at first glance, appear to be at variance with such sources as *U.S. Naval Technical Mission to Japan* [hereafter NavTech] Report A-11 (Washington, D.C.: Govt. Print. Off., 1946), which discusses Japanese naval aviation equipment and carrier design. However, it must be remembered that the goal of the NavTech reports was to gather information after the war to improve the U.S. Navy's own practices. As such, its primary area of interest was documenting late-war Japanese doctrine and equipment, rather than chronicling the development of that doctrine per se. For instance, carriers such as the late-war Unryu-class (which was a derivative of the original *Hiryu* design) did indeed have the ability to perform more operations on the flight deck than their predecessors, and by 1944 Japanese doctrine had evolved to view the flight deck in a different light. However, it must be remembered that these doctrinal changes were the direct result of battle experience (much of it negative) gained early in the war at places such as Midway. As a result, the way Japanese carriers operated in 1942 was different in certain respects from the way they operated in 1944.

4. *ATIG* Report 2, p. 3, and *ATIG* Report 5, p. 3. This was due to the inability of the forced-air ventilation systems used in the hangars to cope with the exhaust from multiple aircraft. *NavTech* Report A-11, p. 9.

5. Aircraft were usually brought to the flight deck via a single elevator for several reasons. Japanese aircraft were segregated by type and stowed in specific portions of both the upper and lower hangars. Fighters were typically stowed forward, dive-bombers amidship, and torpedo bombers aft. Fighter aircraft, requiring shorter runoffs, were sensibly stored forward, where they were also more immediately accessible. Spotting *Akagi*’s antiship strike therefore would have required lifting the torpedo aircraft using the aft elevator, and the Zeros from the fore. Elevator cycles varied depending on raw elevator speed and whether the aircraft was being delivered from the upper or lower hangar. *Akagi* and *Kaga*’s elevators were older, slower models requiring cycles longer than one minute to the lower hangar, and they therefore took longer to perform their evolutions than the newer ships of Carrier Division 2. This was particularly unfortunate in light of *Kaga*’s large torpedo plane squadron.

6. Spotting sometimes required relatively long lateral deck pushes, though the spotting of one aircraft could occur as another was being brought to deck. Nevertheless, a certain amount of "jockeying" was required during such operations.

7. Initial engine start-up was accomplished by a crewman, while air crew were receiving briefings. Detailed information on Japanese take-off procedures was provided by Mr. Nisohachi Hyodo, an expert on Japanese aircraft ordnance, in a letter to Jon Parshall dated 7 February 2001.

8. Hyodo states that Japanese aircraft carriers were equipped with enough bomb carts to re-arm one-third of the carrier’s complement of dive-bombers at a time. Rearmament occurred on the flight deck, immediately prior to engine warm-up. Both the Japanese 242 kg high-explosive bomb and the 250 kg semi-armor-piercing bomb used the same mounting hardware, speeding the process of switching between these weapons considerably. Even so, five to six minutes per plane would be required and would have to be repeated three times to arm the entire force, for
a total of about twenty minutes. To this must be added the time required to load and move the ordnance across the flight deck. Hyodo to Parshall, 7 February and 10 February 2001.

9. This was conducted on the flight deck near blackboards attached to the side of the island.

10. This included an assistant air officer (sho-hikocho) visiting each aircraft in the strike force to ensure that it was running properly. Hyodo to Parshall, 7 February 2001.

11. The fastest Japanese combat launching on record to this point in the war had been for Pearl Harbor, when the carriers launched aircraft at the rate of one every twenty-eight seconds. Gordon Prange, At Dawn We Slept (New York: Viking Press, 1991), p. 490. ATIG Reports 2, p. 2, and 5, p. 3, cite optimal takeoff intervals as being ten seconds.

12. As corroboration of this estimate, one need only look at the operational tempo of the Pearl Harbor Striking Force six months earlier. During that attack, the Japanese were able to spot the second wave attack force of 171 aircraft for launch in fifty-five minutes from when the first attack wave and the formation’s CAP fighters were finished launching at 0620. In this attack, the six Japanese carriers were spotting an average of twenty-eight aircraft per ship, as opposed to the average twenty-one planes Nagumo’s Midway force would have spotted. Using the model we have developed, and adding an additional seven to ten minutes for extra elevator cycles, as well as a longer warm-up time in the early morning air, the figure of fifty-five minutes agrees well with our estimate. Prange, pp. 490–2.

13. Japan’s official war history series, Boeicho Boeikenshujo Senshibu (originally Boeicho Boeikenshujo Senshishitsu, and often referred to in its abbreviated form Senshi sosho [war history]), was published by Asagumo Shimbunsha. The Midway volume, Midowei kaisen [Battle of Midway], published in 1971, states on page 289, “Provided that the strike forces were fully equipped, it would have taken no less than 40 minutes to get them out of the hangar to the deck and then finish preparation for launch.” We are grateful to Nisohachi Hyodo and Takashi Koganemura for their assistance in these matters.

14. Takeoff from a carrier required using full military power. Radial aircraft engines were (and still are) built with very thin cylinder walls to extremely tight tolerances, and they required uniform distribution of heat and lubricant to maintain efficient operation. If not properly warmed up, a radial engine was likely to blow up outright under full power. We are grateful to Eric Bergerud and Clint Bauer for their insights on this crucial issue.

15. It must be remembered that the Japanese operated the aircraft of their carrier groups in a much more integrated fashion than Western navies could at this point in the war. Japanese carrier divisions were trained to combine like-type squadrons into divisional strike teams. Thus, during the initial strike against Midway, the Type 99 dive bombers from Akagi and Kaga formed a single attack unit, likewise the Type 97 attack bombers from Soryu and Hiryu. Consequently, the time requirements of rearming procedures, and certainly deck-spotting activities, were likely to be similar among the carriers of a division, meaning that information concerning a particular carrier can give us clues as to the state of its divisional counterpart as well.

16. Reconstructed from “Mobile Force’s Detailed Battle Report 6,” translated and published in 1947 by the Office of Naval Intelligence as “The Japanese Story of the Battle of Midway, OPNAV P32-100” (often referred to simply as the “Nagumo Report”), pp. 13–20, as well as “Battle Report of Battle of Midway” (extract translation from document 160985B—MC 397.901, U.S. Naval Historical Center, Operational Archives Branch, Washington, D.C.). All times are given in local (Midway) time. We are grateful to both John Lundstrom and Mark Horan for their expertise and assistance in developing a highly detailed and accurate picture of Japanese CAP activities.

17. Ibid. Kaga, as the other big flight deck in Nagumo’s force, carried an equally large CAP burden.

18. Any strike force spotted aft would likely have contained a fighter escort of some sort, requiring Zeros to be brought up from the forward section of the hangars via the forward elevators, thereby obstructing the flight deck for takeoffs in any case.

19. The Japanese Type 0 fighter was not constrained by range or fuel capacity but rather by its cannon ammunition. Each Zero carried
only sixty rounds for each of its two 20 mm cannons, which constituted the main offensive armament necessary to bring down the large American attack aircraft. As a result, Japanese fighters had a tendency to “shoot their bolt” quickly during combat. The importance of cannon ammunition cannot be overestimated—Mark Horan, contributor to *A Glorious Page in Our History* (Missoula, Mont.: Pictorial Histories, 1990), has pointed out to us that casualties among the attacking American squadrons are strongly correlated with whether the Japanese fighters they encountered carried fresh loads of cannon ammunition. Japanese doctrine normally called for two-hour fighter patrols (*ATIG Report 2*, p. 2). During the morning’s air battles, at least seventeen CAP fighters ended their missions after an hour or less, some after as little as twenty-six minutes in the air. This evidence indicates that the Japanese were cycling their fighters frequently in order to keep them fully munitioned.

20. *Akagi*’s strike force was to consist of eighteen Type 97 (Kate) torpedo bombers and three Zero fighters. *Kaga*’s contribution was to be twenty-six Kates and three Zeros, *Soryu*’s eighteen Type 99 (Val) dive-bombers plus three Zeros, and *Hiryu*’s eighteen Vals plus three Zeros.


22. Best, in an interview with John Lundstrom, April 2000. Best stated that during the time of his attack, six to seven aircraft were on the flight deck, and they were clearly Zeros. Furthermore, Best commented that the Zeros were using most of the flight deck for run-off room. As he was attacking, a Zero was in the process of taking off.


24. Given the enclosed nature of Japanese hangars (which amplified explosive effects upon the ship’s structure of internal bomb hits), Japanese carriers with planes in the hangar were in an even more dangerous position than if they had planes on the flight deck. Having no planes aboard when struck was, of course, optimal.

25. Japanese “Mobile Fleet Doctrine,” promulgated 28 March 1944, under “Air Combat,” paragraph 9, states: “When enemy aircraft carriers are discovered at close range the command ‘Send up ‘Q’’ will be given. At this time every ship will quickly send up the airplanes standing ready on deck. The *hikokitai* [carrier air group] will assemble in the air and will fly off to the attack organized in the fixed *hikokitai* [i.e., standing Table of Organization].” Translations of these doctrinal works, recovered from the sunken cruiser *Nachi* in Manila Bay in 1944, are in the personal collection of David Dickson.
THE PERILS OF PAPERLESS
SOME QUESTIONS ABOUT THE LATEST
DEFENSE BUSINESS TREND

Patrick J. Geary

In a document released in November 1997, Defense Reform Initiative: The Business Strategy for Defense in the 21st Century, then–Secretary of Defense William Cohen stated, “To carry out our defense strategy into the 21st century with military forces able to meet the challenges of the new era, there is no alternative to achieving fundamental reform in how the Defense Department conducts business.”¹ One initiative spelled out in the document concerns how Defense Department business practices are related to the management of technical data supporting defense weapons systems.

Citing how recent improvements in information technologies have allowed the business world to conduct numerous operations in a paper-free environment, Secretary Cohen brought attention to the need for the Department of Defense (DoD) to move in the same direction. This department-wide initiative calls for 85 percent of all DoD technical manuals and 80 percent of all technical drawings to become electronically accessible. It is designed to achieve significant benefits: “By integrating paperless technical data management with electronic commerce for business information, DoD will eventually be able to support all major weapons systems in a paperless environment, from the initial design phase through production, operation, and maintenance.”²

If fully implemented, the initiative promises such specific benefits as: a reduction in the cycle time for production contract awards; a reduction in the time to review technical drawings; a reduction in the number of contract data requirements lists needed to conduct business with DoD program offices; and significant cost avoidance.³

Despite these potentially significant benefits, however, the conversion of tens of millions of technical drawings, models, manuals, and manufacturing information into electronic images for easier access has profound implications for the adequate protection of the nation’s most critical and sensitive defense-related information. Two primary concerns are access control to proprietary information and protection of classified information. Without addressing these issues fully, DoD’s “new business strategy”...
might very well have an overall negative impact on U.S. national defense strategy.

PROPRIETARY INFORMATION

Defense-related technical data includes a variety of sensitive (and sometimes classified) information that must receive limited distribution and careful access control. Protection requirements for this information are specified in federal statutes and regulations, as well as directives, instructions, and standards of the Department of Defense, each of the military services, and the Department of Commerce. One type of sensitive data requiring protection from unauthorized disclosure is proprietary information. There are three main concerns with electronically accessible proprietary information: legal liability, labeling and controlling the accuracy of data, and identifying users.

Recent federal court decisions, such as Bernstein v. U.S. Department of State, and subsequent written opinions from legal counsel of the Department of the Navy have focused on the need to protect proprietary data from unauthorized access via the Internet. This legal issue has significant implications for systems that will operate and interface via the Internet. These legal interpretations have specifically stated that failure to properly protect proprietary data could result in violations of federal statutes such as 18 USC 1905, which prohibits the disclosure of proprietary data by the federal government.

The legal issue involving proprietary information is especially complex. Much of what the United States needs to conduct its national defense strategy comes from defense contractors, many of whom rely heavily on their proprietary or trade-secret information to stay in business. They allow the U.S. government access to their proprietary information on the condition it will be protected. If, however, in compliance with DoD’s new business strategy, such information is put in electronic form but then not adequately protected, an unauthorized individual, organization, or company could obtain access to another company’s proprietary data.

Until recently, almost all such data was kept in stand-alone storage facilities, which made it relatively easy to set up access control procedures. However, it is much more difficult to control access to data that is available through the Internet. Therefore, implementing a new DoD business strategy for electronic access poses an increased—and increasing—risk of legal liability to the U.S. government because of inadequate protection of proprietary information.

Also, in all cases where access is limited and distribution strictly controlled, there must be a method for the U.S. government to indicate who may have access. Maintaining control of this data will require labeling. It is not possible to control access to sensitive electronically stored information unless it is labeled in
a way that can be universally recognized and understood. Most of the data described, including proprietary data, is either inadequately labeled or not labeled at all, making it impossible to determine the sensitivity of each item. Providing adequate electronic labels for the voluminous proprietary data will undoubtedly be a time-consuming and costly undertaking.

To date, Congress has provided no funding to ensure proper access labeling of electronically stored proprietary information. Labeling will also be required to indicate whether the given document has been modified or destroyed. Known as “data integrity,” this is one of the most important aspects of network security. Engineers rely on the accuracy of technical drawings for all aspects of deploying and maintaining weapons systems for national defense. Since many people currently have access to this data and could modify it, verifying its authenticity is critical for research, development, testing, evaluation, and production results. Converting current hard-copy data into electronic images will entail the associated, and extremely difficult, requirement to label each file’s (and subfile’s) sensitivity level and unequivocally certify its authenticity.

Electronic data is stored in and transmitted among a large number of repositories, local-area networks, and wide-area networks throughout the United States and several other countries. The objective of the new DoD business strategy initiative is to link all repositories and networks using the Internet to allow faster communication between federal agencies, military departments, and defense contractors. However, the larger the number of users and the more diverse the organizations involved, the more difficult it will be to control the accuracy of data and the identity of authorized users. It has been estimated that soon 1.5 million users will have authorized access to defense-related technical information. These users will be U.S. military personnel, government employees, contractor personnel, and foreign nationals.

Defense contractors are companies of various sizes and organizational structures. Some include many divisions or subsidiaries, while others may be wholly owned, controlled, or influenced by other companies, organizations, or even countries. One division within a company may have authorized access to information for the performance of a specific contract that another division of the same company does not.

This implies that, given the large number of users and all the interconnectivity between repositories, networks, and diverse organizations, converting DoD technical data into electronic images will mean a reduction or loss of the ability...
to confirm the identity, need-to-know, and authorization of each individual wishing access. Also, it will become increasingly difficult to keep proprietary information from being modified, destroyed, or exposed to other kinds of deliberate or unintended unauthorized disclosure, such as hacking.

CLASSIFIED INFORMATION
As the Department of Defense plan unfolds to include classified data among the types of information to be electronically accessible, two unanticipated problems have come to light: how to get the classified data securely to the desktop, and how to store and protect the data once it is there. Fortunately, moving classified information through the Internet via Type 1 encryption is now becoming possible, but desktop storage and user security are still concerns.9

Classified information in storage must be physically separated from other data to ensure its protection. All users handling classified data on the Internet must have independent classified storage and handling capabilities at their desks. Therefore, unless an alternative solution can be found, each user processing digitized classified information on the Internet will need a separate and secure personal computer, or a removable hard drive that is reasonably priced and user-friendly. All of the concerns about processing and protecting corporate proprietary information apply (with even greater stringency) to the processing and protection of classified national security information.

SOLUTION
The solution to the proprietary data problem centers on the labeling issue. Two possibilities come to mind. Congress could appropriate a large amount of money—possibly as much as several hundred million dollars—to create and administer a universally acceptable system of labels. The system would have to administer literally tens of thousands of data categories and access levels. The second solution would be to develop a machine capable of performing the same function.

In conclusion, it is imperative that these problems be resolved before the new defense business strategy is fully implemented. Sometimes the advantages of new technological developments disguise the problems they create. In this case, the problems especially concern data security as part of the overall DoD defensive information operations. For a variety of reasons, data security has, until recently, generally been overlooked as a matter of high priority in the digital world. So far, it appears the United States has been generally fortunate in protecting its data. However, if government funding is not soon forthcoming to accompany the new defense business strategy’s plan for digitization and networked access to vast bodies of sensitive technical data, a dire price might be imposed on individual companies or even national security as a whole.
2. Ibid.
3. Ibid., p. 7.
4. A complete list of the applicable requirements is available upon request.
6. The author learned this while serving as the computer security project manager for a Joint Service Program Office from April 1998 to July 1999.
8. The U.S. Navy and Marine Corps are currently building a joint internet that will accommodate approximately 500,000 users. If one multiplies that number by three to include the U.S. Army and Air Force, the result will be approximately 1.5 million.
9. Probably the best-known secure network now in use is the SIPRNET.
IN MY VIEW

MacARTHUR’S CHRYSLER

Sir:

Your reviewer of my *MacArthur’s War* in the Winter 2001 issue [Dr. Donald Chisholm, of the Naval War College] must have had a bad-hair day. Not a single one of his nit-picking corrections, some of them already altered in the next printing, relate to the thrust of the book (largely ignored in the interest of demonstrating his superior naval expertise), which was that General MacArthur bungled the command of the Korean War by failing to run a hands-on operation and by a pattern of willful and arrogant insubordination.

His technicality that the Japanese minesweepers and crews were not really part of the Imperial Navy obscures the immorality of employing them in a war operation at Wonsan in which at least one ship and crew were casualties.

And he missed at least one more error as crucial to MacArthur’s mismanagement as the rest—now corrected in the paperback reprint. The shiny new vehicle in which the general rode to Haneda Airport as he was exiting Japan was a Chrysler rather than a Cadillac.

STANLEY WEINTRAUB

*Pennsylvania State University*
THEY WOULD HAVE FOUND A WAY

Sir:

The essay by Parshall, Dickson, and Tully in this issue [pages 139–51] was both a critique of my article “The Battle of Midway: Why the Japanese Lost” [Summer 2000, pp. 60–100] and an exposition of their own theory of why the Japanese could not get a strike force launched from their carriers at Midway before those carriers were bombed at 1025 on 4 June. I wish to respond to certain points of their critique of my article and then offer some comments on their theory.

First, I want to commend the authors for producing a most interesting essay. They and I share the belief that the conventional American scenarios of what happened on the Japanese carriers that fateful morning do not make sense, and they as well as I have attempted to fashion more plausible scenarios—based on more recent Japanese sources—to explain why Admiral Nagumo could not get a “grand scale” attack launched. In that endeavor we have come up with very different explanations on certain points, though we agree on others. As one of the purposes of my article was to stimulate critical analysis of the subject, I welcome alternative points of view in the hope that from the clash of ideas a better understanding of what really happened on the Japanese side of that battle will eventually emerge.

Indeed, the authors have made me rethink some of my conclusions. I have even been persuaded to concede one point that affects the timing in my scenario for the operation to rearm the torpedo planes and dive-bombers on the Japanese carriers. I now accept that the second-wave planes were already in the hangars when Nagumo’s order to rearm them was given at 0715. Even had they been spotted on the flight decks soon after the first wave departed for Midway, I now believe that when the first American attack wave from Midway was anticipated (Nagumo knew before 0600 that his carrier force had been spotted), they would have been struck below to free the flight decks for combat air patrol (CAP) activity. The Japanese record of Zeros on CAP being recovered around 0700 on Akagi and Hiryu substantiates this. This has the effect of advancing my rearming schedule for Akagi by the time it would take to get the first few torpedo planes to their arming stations in the hangar. (As I have the rearming process commencing on each plane as it reaches its arming station—about six minutes for the first one, twelve minutes for the sixth—rather than waiting until the entire squadron was lowered, the net advancement would be about ten minutes.)

This does not, however, vitiate my key point, which is: If Nagumo did not receive the Tone 4 sighting report until 0800, instead of 0740 as commonly assumed, the rearming operation would have proceeded twenty minutes longer...
before suspension or reversal than under the conventional scenario. As a result, under my original rearming schedule, all the torpedoes would have been removed by 0800, not only constraining Nagumo’s options at 0830—when a decision had to be made whether or not to launch a strike against the American carrier force—but leaving much more to do after 0830 to restore the torpedoes than if the rearming operation had been reversed at 0745.

Under a schedule advanced by around ten minutes, even more would have been done by 0800—land bombs would probably have been installed on the first chutai (division) of torpedo planes on Akagi. This would have left even more work to be done after 0830 to reinstall the torpedoes—resulting in Nagumo’s 1030 deadline for launch being delayed even longer than under my original schedule. However, this does not negate my supposition that had Nagumo reversed the rearming operation at 0745, as claimed in the conventional scenarios, there probably would have been time to restore the torpedoes in time for a launch to have taken place by 1000.

Having lauded the authors, and even having conceded a substantial point to them, I now turn to some points in their critique of my article that I think are in error.

They dispute my hypothetical assertion that the respotting of the torpedo planes on the flight decks of Akagi and Kaga could have begun at 0920 had they been rearmed with torpedoes (which I contend would have been possible had the rearming operation been reversed at 0745, as has been claimed). They refute this by saying (on p. 145) that this would have been prevented by “the high tempo of CAP operations” during the period commencing at 0920. They say that at 0920 “the Japanese were alert to a constant stream of incoming American strikes. Until the coast was reasonably clear, and his CAP well stocked with fresh fighters to last through the spotting process, Nagumo cannot have been expected to spot his strike.”

I believe this to be overstated. At 0920, Nagumo was aware that one American torpedo bomber squadron was approaching his carrier force (with dive-bombers expected soon to follow.) He had no reason at that time to expect a “constant stream of incoming American strikes.” As for CAP operations during the period after 0920—five Zeros were launched at 0932 and three more at 0945—they were not an impediment to spotting attack planes on the flight deck aft. Although the authors state (on p. 143) that planes would not “usually be spotted aft during fighter takeoff operations” (their emphasis), they give no logical reason why this could not be done if there was a compelling reason to do so. Their note accompanying this statement says, “Any strike force spotted aft would likely have contained a fighter escort of some sort, requiring Zeros to be brought up from the forward section of the hangars via the forward elevators,
thereby obstructing the flight deck for takeoffs in any case.” Only three Zeros from each carrier were to be used for escort. One would think that they could be brought up after the last CAP Zero to be launched prior to 1000 had taken off; that would have been at 0945. There is no reason why they would have had to be spotted at the same time as the first attack planes were spotted aft. Thus, there is no good reason given why Nagumo could not have begun spotting his strike on the flight decks at 0920 had they been ready (which, of course, they were not).

As for my assertion that had (hypothetically) the strike force been ready to launch at 1000, it could have been launched during the fifteen-minute “window” between attacks on the Mobile Force, the authors (on p. 145) counter as follows: “This is wrong on several counts. First, we know that Akagi landed a CAP fighter at 0910 and two more at 0951, meaning that even if there had been strike aircraft on deck at 0920 (which we think unlikely in any case), they had to have been moved back down into the hangar by 0951.” They seem to be saying that a “grand scale” strike would have been aborted—and the strike planes stricken below—in order to land two Zeros on CAP at 0951! I believe that most people would assume that the recovery of those two Zeros for reservicing would have been postponed until the strike was launched. The authors err in assuming that because Zeros on CAP were landed when no strike was spotted, those Zeros would have also been landed in different circumstances—such as when a launch of strike aircraft was imminent.

The authors continue: “Also, Isom clearly does not factor in the immutable time costs associated with spotting and engine warm-up—a ‘fifteen-minute window between attacks’ simply does not suffice.” My hypothetical case assumes that the strike planes would already have been spotted by 1000, and with the engines of all but the last few planes brought up already being warmed up. The engines on those last few planes could be warmed up while the planes in front of them were being launched. In view of this, fifteen minutes would have sufficed to launch the strike had it been ready (as I posit it would have been had the rearming operation been reversed at 0745).

The authors go on to say that I ignore “Nagumo’s own estimate that the strike force would be ready at 1100, although a 1030 takeoff was hoped for, if things went well. Launching at 1000, though, for all the reasons cited above, was never even remotely in the cards, and Nagumo knew it.” Here, they appear to have confused the two rearming scenarios I have been comparing: the one that would have resulted had Nagumo received the Tone 4 sighting report at 0740—as conventionally assumed—as opposed to the rearming schedule that probably actually resulted from his not receiving it until 0800, and reversing the rearming operation after that time.
The 1030 scheduled launch time was a consequence of the actual (and later) reversal of the rearming operation—and of course, a 1000 launch time was “not in the cards.” But my point was that it well might have been had the rearming operation been reversed at 0745. (Incidentally, the authors have misread Nagumo’s statement in his official report. Nagumo states, on page 7 of that report, that he was advised that the torpedo-equipped attack planes in Carrier Division 1 (Akagi and Kaga) would be ready for takeoff at 1030, and that the torpedo planes in Carrier Division 2 (Hiryu and Soryu)—which had returned from the Midway strike—would be ready by 1030–1100. These latter planes would be in a strike group separate from the one he hoped to launch at 1030, which included the torpedo planes only of Carrier Division 1.)

Having lambasted the authors on the preceding points, I now want to concede another point to them. In their next paragraph (on p. 145) they state: “Furthermore, the assertion that two-thirds of Akagi’s torpedo planes were on deck at 1000 is clearly wrong.” Here, I believe the authors are correct. My reconstruction of the actual rearming schedule was premised on the second-wave planes being lowered to the hangars after Nagumo’s rearming order was issued at 0715. As discussed earlier, I now accept that they were already in the hangars at 0715. As this would have advanced the rearming schedule by at least ten minutes—requiring even more work to be done to reverse it after 0830—it would have resulted in fewer torpedo planes being ready for respotting by 1000. Whether or not any of Akagi’s torpedo planes actually got up to the flight deck before it was bombed is, despite Senshi Sosho’s claim, still debatable (though I am now convinced that none of Kaga’s were on the flight deck).

Now for a few comments on the authors’ theory of why Nagumo could not get a “grand scale” strike launched before his carriers were bombed at 1025. They contend that beginning with several minutes before 0700 and running until 1030, the need to keep the flight decks free for CAP activity prevented the spotting of strike planes for a launch at all times during that (three-and-a-half-hour) period. Such a launch was precluded, they say, because it would require forty minutes to raise and spot a squadron of torpedo planes or dive-bombers, plus additional time to warm up the engines and make the launch—adding up to nearly an hour. (It is said that the Zeros on CAP could not be deprived of servicing or reinforcements for that long.) They conclude that the inability of Nagumo to launch a strike “hinged neither on whether Nagumo received Tone 4’s message at 0740 or at 0800, nor on how quickly the armorers in the Japanese hangars could do their work.”

The implications of this theory are astonishing. According to its logic, Nagumo would not have been able to launch a grand-scale strike against the American carrier force even had he not rearmed his second-wave torpedo planes and dive-bombers for a second strike on Midway, and even if he had received the
Tone 4 sighting report immediately after it was sent at 0728, and even if a more thorough search effort had discovered the American carriers at 0700. The blunders committed in rearming, search, and communications operations, which have been blamed by Japanese as well as American historians for the debacle that befell Nagumo, were—we are told—irrelevant. Even if Nagumo’s torpedo planes had been properly armed at 0700, the authors contend, they could not have been launched, because they were in the hangars at that time and could not have been raised and spotted on the flight decks before 1030—until the American attacks were over.

This theory, which completely fineses my analysis and that of many others, is in my view simplistic. While the authors are to be commended for bringing to light complications in Japanese carrier operations caused by CAP activity—complications underappreciated by commentators on the battle, including me—they have applied certain elements of their theory much too rigidly.

First, regarding elevator operations: They contend that forty minutes was required to raise from the hangar and spot on the flight deck a squadron of planes. Although it could take forty minutes to raise and spot a squadron of torpedo planes on Akagi, this was true only if one elevator (the aft one) was used. Only the aft elevator could be used when a full air group was in the hangar, as dive-bombers would block the use of the midship elevator for raising torpedo planes. (This would have been the case after 0900, when the Midway strike dive-bombers, having returned, were stowed in the hangar.)

However, several Japanese veterans of the battle whom I interviewed stated that if the dive-bombers were already aloft—as Akagi’s were after the Midway strike force departed—the middle elevator could also be used to raise (or lower) torpedo planes in an emergency. Likewise, when Hiryu’s and Soryu’s torpedo planes were aloft (such as before 0900), the aft elevator on those carriers could be used to raise dive-bombers. (Unlike Akagi’s aft elevator, the ones on those carriers were large enough to accommodate dive-bombers.) This would reduce the elevator time by almost half. Also, the elevators on the newer Hiryu and Soryu were faster than those on Akagi and Kaga. Dive-bombers on those carriers could be raised and spotted in less than forty minutes even if only one elevator was used. Thus, it did not always, on all carriers, take forty minutes to raise and spot a squadron of bombers.

This faster elevator operation for the dive-bombers on Hiryu and Soryu is implied by Senshi Sosho (the official Japanese history of the battle, which the authors appear to accord a great deal of credibility). It states (on pages 289–90) that the dive-bombers could have been launched very soon after 0830 (and Minoru Genda, Nagumo’s air officer, states that at least some of those bombers were already on the flight decks at around 0830). But the authors, rigidly applying what they believe to be Japanese carrier doctrine, state (on p. 146) that
the “dive bombers were all in their hangars at the time and would have taken another forty minutes to put in action, even if they had been rearmed.”

Yet Genda, Ryunosuke Kusaka, and Tamon Yamaguchi thought that the dive-bombers could be launched soon after 0830. Even if they were in the hangars at 0815, those Japanese officers apparently relied upon a fairly quick raising of them to the flight deck, utilizing two high-speed elevators on each of the carriers. For the authors to insist that only one elevator could be used for each squadron of bombers and that it would invariably take forty minutes to raise and spot them on the flight deck regardless of the circumstances and gravity of the emergency is, in my opinion at least, much too extreme.

Second, regarding CAP operations: Although they clearly constrained deck-spotting operations of the bombers, the constraints were not as absolute as the authors maintain. I have already pointed out that CAP takeoffs did not prevent the spotting of bombers aft. Landings of Zeros for resericing did require a free flight deck aft, but there was much more flexibility than the authors allow. For example, they state that as Hiryu recovered CAP at 0840 “even if strike planes had begun to be promptly brought up on Hiryu after the American B-17s departed at 0815, they would have had to be stowed below again by 0840.” Again, I believe most people would assume that landing of the Zeros would be postponed until the strike had been launched. To suggest that a strike ready to go would be aborted in order to land some Zeros on CAP seems much too dogmatic.

Thus, we are told that even the option of a launch of dive-bombers alone at 0830—an option that Nagumo has been roundly criticized for not taking up—was in fact actually precluded by Japanese carrier doctrine relating to elevator operations and CAP. Likewise, we are told that this carrier doctrine precluded the spotting of a strike between 0920 and 1000 and its launch during the “window” between 1000 and 1015, even had one been ready.

While an underappreciation of the constraints that CAP operations placed on strike plane operations may have been the greatest weakness in my analysis, it seems to me that dogmatism by the authors regarding Japanese carrier doctrine is the greatest weakness in their essay. I still believe that Nagumo and Genda would have found a way to spot and launch a strike before 1025 had one been ready. Therefore, I still believe it relevant to inquire why one never got ready.

DALLAS WOODBURY ISOM

London, United Kingdom
WHERE WILL PRESIDENTIAL AUTOCRACY TAKE RUSSIA?

Sergei Khrushchev


This book describes the Russian presidency between 1990 and 1996, its society, and its politics in the “Second Russian Republic” between 1993 and 1996. Examining one of the most dramatic periods in the history of Russia, Nichols begins with Mikhail Gorbachev’s attempt to pull down the old authoritarian system and to push Russia onto a democratic path of development.

In the first two sections, the author—a Naval War College professor—briefly recounts the events of the years 1985–1991, offering his own interpretation of Gorbachev’s failure. He poses a question: What kind of democracy, parliamentary or presidential, suits Russia better? This issue has never before been discussed in this way, even though it is a most urgent topic with respect to what has been happening, and is happening now, in Russia. Nichols concludes that presidential democracy is preferable, maintaining that

“presidentialism in Russia is not a ‘mistake,’ an experiment, or an authoritarian hoodwinking of the public, but rather a deliberate act, a compromise among elites who, like the public that elected them, see it as the system most likely to protect all of them from each other.”

Furthermore, the author finds confirmation of his thinking in the chaotic parliamentary democracy of Gorbachev’s time, when both politicians and the people, unaccustomed to liberty, fell upon each other
with mutual accusations, paralyzed the office of the president, and disoriented the whole of society, all of which rapidly led to the disintegration of the economy, inflation, and loss of central control over vast regions—culminating in December 1991 in the breakup of the country and the departure of Gorbachev from the political scene.

As a counterbalance to Gorbachev, Nichols advances Boris Yeltsin, who was elected in June 1991 as the president of Russia. Yeltsin began a bitter struggle with parliament and parliamentarism. He struggled for the establishment of strong presidential power—so strong that soon his democratic-reformist appointees (especially Boris Nemtsov) almost openly called the president a tsar.

Nichols is absolutely right. Between the uncircumscribed freedom of parliamentarism, which in Russia’s case was accompanied by anarchy, and a presidentialism that resembles monarchism, the latter is preferable. In the former case, the response to anarchy would be an even harsher dictatorship, but in the latter case there could be hope that the country would pass over the reefs of a transitional period and gradually enter the mainstream of normal democratic development.

Nichols argues his case quite persuasively, dividing Yeltsin’s presidency into the “First Republic” (up to the shelling of parliament by tanks in October 1993) and the “Second Republic,” when Yeltsin carved out for himself a presidential republic, in which the Duma became in large measure a deliberative organ, as during the rule of the last Russian emperor, Nicholas II. In the author’s opinion, this turnaround allowed Russia to overcome the political crisis of 1993 and offered the chance for peaceful development. This fairly detailed account of the events of those years closes with the election of Yeltsin for a second term as president of Russia in June 1996.

Where will presidential autocracy take Russia? Nichols poses the question without answering it, for no answer exists. However, the book obliges the reader to consider deeply the complexity of introducing democracy into an undemocratic society, and the vicissitudes of that process. Unfortunately, in explaining the difficulty of the democratic transformation in Russia, Nichols makes the usual mistake of Western studies of describing all obstacles as proceeding from the totalitarian Soviet epoch, the atomization of Soviet society, etc. In fact, everything is much more complicated. The Soviet period was undoubtedly totalitarian, notwithstanding its Marxist ideological dogma, little different from the preceding centuries of Russian monarchical absolutism. Whereas the West, especially the United States, grew out of the Roman tradition of respect for written law, and therefore for constitutions, or basic laws, Russian political culture matured with the Byzantine emphasis on the Will as something higher than the Law. Seventy years of Soviet rule did not change the Russians; they reinterpreted
Marxism, and Western ideology, in their own fashion, just as they are now trying to reinterpret Western democracy. Therefore, Russia’s progress toward a normal democratic government will be more painful than Nichols represents.

Having given a history of the Russian presidency, the author unfortunately limits himself to a superficial account of the myriad interparty confrontations. Within Russia all the “party” intrigues appear to be merely reflections of a struggle for national power among oligarchic-criminal groups that emerged as a result of fraudulent privatization—groups that control political parties, power structures within the government and the administration, the press, and television. Regrettably, this key aspect of political life in both the First and, especially, the Second Republics is completely absent from the book, a fact that substantially lessens its value for understanding what is and has been going on in Russia. Neither does Nichols in his analysis hesitate to indulge an easy division of historical players into good and bad.

For example, he paints the speaker of the parliament in the First Republic, Ruslan Khasbulatov, and Vice President Aleksandr Rutskoi in exclusively dark tones. Such an interpretation of events simplifies the author’s exposition but distorts the historical picture. In the parliament-president conflict of 1992–93 both sides were to blame—and perhaps Yeltsin, with his pathological striving for personal power, more than Khasbulatov. Undoubtedly, the Yeltsin-Khasbulatov struggle had to result in the elimination of one of them from the political arena. Nichols seems to welcome Yeltsin’s victory, but The Russian Presidency would doubtless have benefited had the author turned his attention to the forces that the two players stood for. It would have been interesting to examine the possible results had Yeltsin not signed the unconstitutional Decree 1400 that dissolved parliament and the constitutional court, and suspended the operation of the constitution itself.

In addition, not to the author’s credit are several political clichés that have been transferred to this serious, historical work from the pages of periodicals. Thus Vladimir Zhirinovskiy and his Liberal-Democratic Party are presented by the author as a demonic, fascist opposition force. Yet Nichols declares it to be well known that Zhirinovskiy and his party have always been controlled by the government and the president, voting in parliament as ordered by the Kremlin, and that its extremist-hooligan rhetoric serves a single goal, to divert the populace from the actual oppositional and protofascist movements, like the Russian National Unity Party.

Also questionable is Nichols’s contradistinction in the last chapter between Boris Yeltsin and the Belorussian president, Aleksandr Lukashenko, as a civilized and powerful president versus an abominable dictator, respectively. In such a scientific-historical work, such propagandistic methods used without concrete
evidence are inadmissible and reduce confidence in the author. If one is to be objective, Lukashenko is the very image of Yeltsin. In 1994, he repeated everything that Yeltsin did in 1993 but without having tanks shoot at the Belorussian parliament. The emergence of president-autocrats has been a phenomenon in the post-Soviet era; Yeltsin and Lukashenko, far from being exceptions, require separate, serious analyses along the lines established by this book.
BOOK REVIEWS

THE CONSEQUENCES OF NUCLEAR PROLIFERATION


The Coming Crisis is a series of essays by noted scholars in the field of national security affairs examining the effects of continued nuclear weapons proliferation and the potential for regional nuclear crises. While one can argue that use of nuclear weapons by a rogue state today is more likely than it was during the Cold War, discussion of what the United States could and should do if deterrence fails has been noticeably absent in recent years. The authors revive this discussion and, in doing so, contribute significantly to the study of proliferation.

The first part of the book addresses the underlying pressures that cause states to consider acquiring and potentially employing nuclear weapons. It makes the case that there are many reasons why a state may develop nuclear weapons, reasons that may have only a peripheral relationship to security issues. A highlight of this section is Caroline F. Ziemke’s essay on the strategic personality of Iran. She asserts that the behavior of a “rogue state” often has roots in national myth. These nations may not be as unpredictable and as roguish as we might surmise at first glance.

The second part of the book consists of five essays that examine how nuclear crises might develop between the United States and a regional nuclear power, and what the consequences might be. Stephen Peter Rosen and Stephen M. Walt each examine the impact of proliferation on alliances and coalitions, but they come to contradictory conclusions. Barry R. Posen conducts a hypothetical analysis of a Gulf War with a nuclear-armed Iraq. His conclusion is that, faced with such a crisis, the United States should not and probably would not hesitate to intervene to defend vital interests.

Two essays focus on what the United States might do following a first use by a regional aggressor. George H. Quester argues that a U.S. response in such a case would be driven more by American norms of law enforcement than by Cold War theory. Quester believes that criminals are punished for four distinct reasons: to disarm, to make an example, to impose revenge, and to reform. He states that once a regime with a modest nuclear capability has used nuclear weapons, it will be seen as “too dangerous to live
with... but not too dangerous to defeat.”

Similarly, Brad Roberts discusses regional nuclear war termination, arguing that the United States would have not only to address the immediate problems presented by the war but also to ensure that longer-term U.S. interests were served by “winning the peace” that follows. The United States has to avoid being perceived as a “nuclear bully” whose power must be counterbalanced, but neither can it come off as a “nuclear wimp,” unwilling to confront an aggressor. Instead, the course of action chosen must show the United States to be a responsible and just steward of the international good.

In the concluding chapter, Victor A. Utgoff contends that in response to a regional nuclear threat the United States would likely be far more resolute than others have suggested and would likely respond in kind to a first use of nuclear weapons by an aggressor. He concludes with a number of policy implications.

The fundamental premise of this book is that sooner or later the proliferation of nuclear weapons is going to lead to a confrontation between the United States and a nuclear-armed state. While there are many points of disagreement between the authors, all concur that such a confrontation will be a seminal event and will define not only the role of nuclear weapons but also that of the world’s only superpower in the post–Cold War era. All students of national security policy owe it to themselves to consider the policy implications of this premise. The Coming Crisis will be valuable to them.

JON GREENE
Commander, U.S. Navy


The title says it all. This book is a compilation of empirical and analytical data on the strategic evolution of nuclear, biological, and chemical (NBC) agents and weapons in the twenty-first century. A central theme of the book is how new regional players (states and nonstate actors) are likely to convert prevalent conventional military doctrine and training into nonconventional means of warfare. The book is very ambitious in its scope; it attempts—overall, successfully—to address systematically conceptual problems in the integration of such weapons into the military infrastructure, delivery systems, command and control procedures, and war plans. More importantly, the editors and the authors of the various case studies utilize a theoretical framework to explain and predict future trends of behaviors, intentions, and capabilities among very diverse players. Realism and neorealism, organizational theory, and culture are used to flesh out these unique differences in approach as well as in the implementation of NBC programs and doctrines.

Except for the conclusion and the chapter on terrorism, the chapters are case studies, focusing on Iraq, Iran, Israel, India, Pakistan, and North Korea. The authors are specialists who devote a great deal of effort to describing the relationship between strategy and policy, on one hand, and between national security and national military strategy, on the other. The result is a complex web of relationships, behavioral manifestations, and decision-making processes involving an amalgam
of scientific, bureaucratic, and military institutions and forces. In the chapter on Iran, for example, Gregory Giles eloquently argues that Iran was reluctant on moral and religious grounds to use chemical weapons during the first few years of the Iran-Iraq War (1980–87) but that its policy changed abruptly as a result of rising Iranian casualties and fear of Iraqi chemical-warfare preponderance. Hence, realism became key to explaining the Iranian NBC doctrinal shift after 1987. Although Iran ratified the Chemical Weapons Convention within a few months of its coming into force, Iran has opted to pursue a clandestine NBC program. Such weapons are the subject of intense debate within the increasingly factionalized, institutionalized, and secularized Iranian political elite today. This has given rise to “multiple actors playing roles in a key strategic program[,] . . . [ensuring] that there will be continued bureaucratic competition for resources, missions, and influence.” More significantly, such competition has far-reaching political, economic, and military implications, associated primarily with command and control mechanisms. Israel (a chapter by Avner Cohen) and India (by Waheguru Pal Singh Sidhu) also utilize NBC secrecy and ambiguity to enhance their conventional deterrence capabilities—Israel to keep its Arab adversaries off balance and to avoid American nonproliferation pressure, and India to keep China, not Pakistan, in check. Pakistan (Zafar Iqbal Cheema), however, apparently sees the development of its NBC program as a necessity—not a choice—because in the “absence of conventional security alternatives and nuclear security guarantees . . . [such] weapons were viewed as a necessary counter [to] a perceived threat from India.”

The chapter by Jessica Stern analyzes the dynamics of terrorism in the twenty-first century. She argues that the potential for nonstate actors to acquire, develop, deploy, and use NBC weapons is growing. Stern may be correct in her concern. Yet, although there is a precedent, in that terrorist groups such as the Japanese Aum Shinrikyo have used such devices, there is no hard empirical data to support a sustained argument that terrorists will be going the NBC route, at least in the near term. Terrorism has become complex indeed; acquiring, assembling, deploying, and using NBC agents does not mean that the selected device will be workable or effective. Moreover, the cost of embarking on such a program for terrorist causes will almost certainly outweigh the benefits. Terrorists, at least for now, will continue to opt for conventional weapons, albeit at more sophisticated and lethal levels. There is evidence, however, of more credible linkages between terrorism and technology, and between terrorism and international finance.

Ultimately, one cannot escape the fact that among the newly emerging NBC players there is a diversity of doctrines and command structures. This will admittedly make it harder to predict possible political and military outcomes; more significantly, it means that the NBC genie cannot be put back in the bottle. Will the current U.S. debate on the Strategic Missile Defense Initiative exacerbate this already volatile situation?

If there is one criticism to make of this book, it is that it sometimes suffers from a lack of consistency in terms of units of analysis under examination (nuclear, biological, and chemical agents vary remarkably in scale of effects, timing, etc.); some essays weigh more heavily on one agent at the expense of the others.
Comparative analysis should instead generalize, with rigor, about similarities and differences with respect to common phenomena. This book is, however, a valuable addition to the complex body of literature on strategy, national security, and comparative political and military dynamics.

K. A. Beyoglu
Marine Corps Command and Staff College


Although we lived with the dangerous specter of nuclear attack for more than fifty years during the Cold War, concerns about the proliferation of weapons of mass destruction (WMD) have virtually exploded into our consciousness in the past decade. Since the demise of the former Soviet Union—once referred to as our “malefactor partner in the concept of mutually assured destruction”—our fears seem to focus far less on the threat of nuclear holocaust, and more on the threat of attack by chemical or biological agents. The logical point of departure for this shift in focus seems to be the Persian Gulf War, when the world learned of a rogue nation seemingly bent on proliferating these weapons of mass terror.

In this book, Albert J. Mauroni attempts a historical recounting of U.S. efforts to deal with chemical and biological warfare agents on the modern battlefield. Mauroni, a former U.S. Army Chemical Corps officer who currently works as a management consultant specializing in Department of Defense chemical and biological defense programs, provides a detailed look at what was essentially a “cold start” go-to-war effort on behalf of the U.S. armed forces. The consistent premise throughout this work is that no one in the Department of Defense (with the exception of the Army’s Chemical Corps) was even remotely prepared for an encounter with chemical or biological agents as it readied for war with Iraq. Convinced at the onset of Operation DESERT SHIELD that Saddam Hussein would indeed use WMD against U.S. and coalition forces, the Pentagon began what Mauroni describes as a “mad scramble” to train and equip U.S. forces to operate in the presence of WMD agents. He reviews the preparation to defend against exposure to these agents, and assesses U.S. efforts to protect its forces against a highly lethal asymmetrical threat. In addition, Mauronidevotes a chapter to the issue of “Gulf War illness,” providing a fairly meticulous and forthright discussion of this controversial subject. He concludes with substantive recommendations on where the future focus of U.S. efforts to deal with the burgeoning threat of chemical and biological agents should lie. At a minimum, Mauroni’s work at dissecting the policies and decisions of the Gulf War is important if only as a lesson that the United States must never again be so fundamentally ill prepared to operate in the asymmetrical environment.

There are criticisms to be made of this book, however. At the surface level, Mauroni uses far too many acronyms for the book to be easily decipherable for the non-Army (and especially nonmilitary) reader. Although he includes a list of abbreviations at the beginning to assist with the veritable “acronym soup” of abbreviations, it becomes confusing and tiresome to refer back constantly to a
glossary to understand what one is reading.

Additionally, Mauroni’s use of the term “chemical-biological” can lead one who is uneducated in the specific characteristics of chemical and biological agents to believe that there is no readily discernable difference between the two types of WMD agents. In reality, there is nothing farther from the truth. Chemical and biological agents are so different in their properties and potential effects on the human physiology that discussions about countering or mitigating their effects should remain separate. By consistently lumping them together, Mauroni gives the reader the impression that the measures taken for defense and consequence management against chemical-agent exposure will be essentially the same as those for coping with a biological threat.

Of greater concern, however, are statements made by the author in the first chapter. He describes his purpose in writing the book: “Only if CB weapons were used on civilians and population centers would they truly be ‘weapons of mass destruction.’ On the military battlefield, these weapons, shorn of the ridiculous air of menace given to them by politicians and the media, are merely another tactical-operational factor like enemy air attacks or unforeseen terrorist attacks; military forces can and do take steps to minimize the effects of chemical-biological contamination. If a military force invests a small amount of time and funds in planning, defensive equipment and training, the immediate threat of mass casualties is avoided, and chemical-biological weapons become merely ‘weapons of mass disruption’ [his italics] instead of destruction.”

I find these comments both naïve and dangerously out of touch with the reality of WMD agents, and certainly contradictory to U.S. efforts at counter-proliferation throughout the Department of Defense. Although we have taken steps to deal with the possibility of chemical exposures among our operational forces (which should be construed as tactical events in the scope and scale of their effects), calling a biological agent a “weapon of mass disruption” ignores its potential for strategic impact. The World Health Organization (which currently offers the most widely accepted casualty estimates for biological agents) predicts that the United States could incur more than 250,000 casualties in a targeted population of 500,000 from only fifty kilograms of weaponized anthrax; such an event could hardly be usefully characterized as a “disruption.”

There are other places where the reader may take exception to Mauroni’s statements—most notably, his comment in chapter 3 that in 1990 the “official U.S. policy was to reduce the likelihood of enemy chemical weapons use by threatening retaliation with similar munitions.” The United States never considered the use of chemical weapons in the Gulf War, since it had long before decided not to use chemical weapons as retaliation in kind. The author’s footnote in chapter 4 regarding the requirement for a company of bakers to augment a medical unit is flatly derogatory to the medical professionals who did so much to ensure that health-protection measures were in place during the Gulf War. His claim that the medics were ill trained and ill prepared to deal with contaminated casualties since “these practical issues had never surfaced in the minds of the medical community” is patently false. Several hundred volumes published since the First World War deal with the medical
handling of chemically contaminated casualties.

In the end, this work comes off as not much more than a “hoo ah” for the Army Chemical Corps, who are billed as having redeemed the Department of Defense’s collective ineptitude with respect to chemical or biological attack. While Mauroni does offer an accurate overall accounting of the Army Chemical Corps’s efforts to deal with the asymmetrical threat of chemical and biological agents on the battlefield, he gives little more than a passing nod to the overall efforts of the other services and their collective attempts to counter or mitigate this omnipresent threat. Readers familiar with the subject of WMD should be cautioned that there is much with which to find exception in this work. Readers unfamiliar with the subject should be careful not to conclude that the capabilities of the Department of Defense are so uniformly one-sided.

PIETRO D. MARCHELLA
Lieutenant Commander
Medical Service Corps, U.S. Navy


Is there need for yet another book on the role of the military in the information age? To judge by this volume, a collection of essays published under RAND’s Project Air Force, the answer is yes—but this may be twice the book we need. In this case, more than enough is not necessarily better. The Changing Role of Information in Warfare is part of RAND’s Strategic Appraisal series, and it primarily addresses the effects of information technology on American military planning and operations. The fifteen chapters provide a useful review of the dangers and opportunities that information technology presents to U.S. military forces. While originally intended for the Air Force, the work should interest a wider professional audience, especially because it includes a broad spectrum of views, ranging from techno-optimists to info-war pessimists.

The editors are well regarded authorities: Zalmay Khalilzad is a former assistant deputy under secretary of defense for policy planning, and John White is a former deputy secretary of defense. Many of the articles were written by well-known writers on strategy and information warfare, and the foreword is by Andrew W. Marshall, Director of Net Assessment, Office of the Secretary of Defense; he is considered by many to be among the foremost thinkers in the U.S. government on future threats and strategies.

However, roughly half the articles cover ground familiar to anyone who has been following the discussion in recent years about the importance of information in warfare and the dangers of computer warfare. These chapters might be useful, for example, to someone looking for a review of the various ways computer hackers can disrupt military operations. But because so much has been written on this constantly changing topic, the more technical chapters do not cover much new territory and are already slightly outdated.

The chapter on information-age terrorism, for instance, warns that future terrorist attacks may take the form of “cybotage” aimed at information infrastructure. This may be true, although it hardly is a new idea; moreover, so far in the information age, old-fashioned terrorism remains dominant, as the attack on the USS Cole reminded us.
Similarly, the chapter on U.S. strategic vulnerabilities discusses the tentative steps being taken at the national level to deal with the information warfare threat—but it has been dated by more recent efforts at information warfare defense, including the Clinton administration’s National Plan for Information Systems Protection, published after this book went to press.

Luckily, only about half the book discusses the familiar territory of information systems and technology. The second half examines many of the broader questions involved in how the U.S. military is adapting to the information age.

One of the best chapters is “The American Military Enterprise in the Information Age.” The late Carl H. Builder argues that the most important effect of the information revolution may not be the application of technology to existing missions but the need for the military to adapt to, and find, new and different missions.

For Builder, it appears that the American military’s “enterprise”—its primary purposeful activity—is no longer (if it ever was) to “fight and win our nation’s wars.” Deterrence and forward defense will not play the central roles they did in twentieth-century conflicts, he speculates, so the military may find itself reduced to providing constabulary and expeditionary capabilities, while keeping the military arts and sciences alive for the future.

Jeremy Shapiro takes a skeptical approach to the entire concept of an “information revolution.” He argues that the information age is not producing the sort of wholesale change we would expect from a revolution, either in military affairs or in society at large. He cites the work of Stephen Biddle and others who have described the “productivity paradox”—the idea that the outlays for information technology have not as yet led to the increases in productivity that would be expected in a social and economic revolution.

If the change is not revolutionary, Shapiro argues, the U.S. military should not hasten to make radical organizational or other changes. He quotes approvingly Eliot Cohen’s observation that the creation of a corps of “information warriors” today might make as little sense as would the creation of a corps of internal-combustion warriors in the last century.

A chapter by Stephen T. Hosmer offers a welcome look at psychological operations (PSYOP), an important aspect of information warfare that is usually neglected by all but the U.S. Army. Army PSYOP advocates may not like what they read here, however. Hosmer argues that although psychological effects are indeed vital, history shows that actual PSYOP efforts are not nearly as effective in reducing the enemy’s will to fight as are well-planned combat operations. Standard measures of psychological warfare effectiveness, such as the numbers of enemy surrenders and desertions, do not correlate directly with the intensity or quality of PSYOP efforts but closely reflect the nature of combat operations.

Military commanders can best produce catastrophic disintegration of enemy resistance, Hosmer writes, not through leaflets and loudspeaker broadcasts but by sustained, weeks-long air and artillery attacks combined with deliberate efforts to deprive the enemy of food, and ultimately with ground operations aimed at exploiting the enemy’s weakened morale. His advice may appear obvious, but it suggests that commanders may be paying too much attention to technical PSYOP efforts and not enough to the psychological
effects of combat operations. In addition, his research seems to suggest that sustained, well-planned strikes may be more important than the sudden mass attacks designed to produce “shock and awe” that are heralded by many network-centric-warfare advocates.

Additional useful chapters review the ethical considerations arising in information warfare and examine whether or not such mechanisms as arms control and export regimes can apply to information warfare technologies. An article by Francis Fukuyama and Abram N. Shulsky reviews the lessons (familiar to a Naval War College audience) that the military can learn from business in adapting to the information age.

One minor complaint—the book does not offer biographical sketches of the contributors. A few pages devoted to that information would be more useful than the largely unnecessary listing of abbreviations and acronyms. Overall, this collection is useful, but a better introduction to many of these concepts is found in an earlier RAND work by John Arquilla and David Ronfeldt, *In Athena’s Camp: Preparing for Conflict in the Information Age* (1997) [reviewed in the Spring 1999 issue].

ERI C J. DAHL
Commander, U.S. Navy
Naval War College


Since 1989, the U.S. military has been involved in a number of intrastate conflicts integrally related to ethnicity. These ethnic conflicts have been devastating to those involved; the conflicts contributed to regional destabilization; and they have been assumed to breed international terrorism. Most saliently, they have virtually destroyed the hope of peace benefits that were predicted to accrue at the end of the Cold War.

The intelligence community was tasked by the State and Defense Departments to provide explanations for ethnic conflict. Indications-and-warning systems were to be developed and used to alert policy and military decision makers to impending crises. It was assumed that good analysis and prediction would contribute to policies and practices designed to prevent, manage, or contain ethnic conflict and thereby minimize damage to international peace and stability. A number of studies were conducted internally or were outsourced. The task was apparently, but deceptively, simple—produce a predictive model of ethnic conflict. The criterion for a successful model was equally simple—did it work? That is, did the model provide more information of a critical nature than could be provided by country experts, and was it available in a timely fashion?

*Identifying Potential Ethnic Conflict* is the public report of research sponsored by the deputy chief of staff for intelligence of the U.S. Army. It was produced by a group at the top level at RAND Corporation in Santa Monica, California.

The stated purpose of the project was to help the intelligence community order its thinking about the logic and dynamics of ethnic conflict and to systematize information-collection requirements. The authors did not provide a comprehensive explanation of ethnic conflict but attempted to answer the questions of how ethnic mobilization occurs and under what conditions it leads to violence.
The research was based on the assumption that reliance on static indicators and simple statistical correlation (as found in many other models) did not adequately account for change. Change was assumed to be the political mobilization of “ethnic factors used to aggregate and articulate group grievances.” Political mobilization was assumed to be found in changes between and among groups in economic, political, and social spheres of life activities. In other words, they were looking for specific changes in the relationship between group and state that could signal future conflict.

Three stages were identified: Potential for Strife, Transformation from Potential to Likely Strife, and From Likely to Actual Strife. The potential for strife was identified in processes associated with closure, that is, the reification of group boundaries, the strengthening of “us-them” thinking. The transformation from potential to likely strife—critical to the conceptual framework—was found in the mobilization of ethnic-political identity influencing the balance of power between a group and the state. This transformation was assumed to be found in factors associated with emerging leadership, mobilization of group resources, and a series of “tipping” events (similar to the “trigger mechanisms” found in other studies). Change from likely to actual strife occurred through an interactive strategic bargaining process; this was portrayed by assessing the group preferences and capabilities of the state and the contending ethnic groups, presented in tabular form. The model was then applied to Yugoslavia, South Africa, Ethiopia, and Saudi Arabia.

The best part of the model is its focus on change and process. The authors probably have that right—most ethnic groups, under most conditions, live in relative peace and harmony; changes in relationships are generally associated with competition and conflict under conditions of relative scarcity. But, however good this assumption, the model was not adequately developed to test the hypothesis on a variety of cases. The variables, as discussed in this publication, were very general; they need further specification and elaboration. The four case studies, which purported to apply the model, were written by situation experts. They are very informative and make good reading. Nevertheless, the writing seemed to reflect the authors’ expertise and perspectives as much as any application of the model.

If, then, the goal was to provide a model that could apply statistical methodology to comparative data and thereby contribute to predicting future ethnic conflicts, it was not accomplished. Most of the conclusions seemed drawn from the analyses provided by RAND’s experts and not produced by the operationalization of the concepts or an application of the model to the four cases. Perhaps because this was a public document, the actual data lies elsewhere and the model has greater specificity and applicability than appear in this short text.

The critical test of any model is whether it works—whether it provides more predictive power than an informed observer. In this case, it is hard to say, because as the authors note, “the model needs further specification and elaboration.” This will not be the final book on ethnic conflict.

As the consultants’ favorite saying goes: “Progress has been made, but further research is necessary.”

PAULETTA OTIS

University of Southern Colorado

Michael Swaine, author of the outstanding *The Military & Political Succession in China* (1992), and his fellow RAND analyst Ashley Tellis have written a very good book about Chinese security in both historical and future perspective. It will certainly be of interest to the policy community, as it should be to all who work on Asian security.

Swaine and Tellis define grand strategy as a country’s “basic approach to political-military security.” China’s grand strategy seeks to preserve domestic order, defend against external threats, and eventually attain “geopolitical influence as a major, and perhaps, primary state.” These bland objectives become vibrant when viewed in historical context—many centuries ago, with strong leadership and domestic order, China dominated the region not only, or even primarily, militarily but also in cultural, political, and economic ways that elicited deference and reduced the need for military capability. Subsequent periodic weak leadership and domestic disorder reduced China’s ability to resist persistent threats from beyond its long, vulnerable border and diminished its political strength, so China has been united as a single entity under Chinese rule for only about half of the last 1,800 years. Today, China is pursuing a pragmatic strategy that emphasizes “the primacy of internal economic growth and stability, the nurturing of amicable international relations, the relative restraint in the use of force combined with increasing efforts to create a more modern military, and the continued search for asymmetric gains internationally” (e.g., entrance into the World Trade Organization as a developing, rather than developed, economy). Assuming that no catastrophic revisions of this strategy are forced on China, Swaine and Tellis expect it to continue through 2015–2020, which they identify as the minimal time frame in which the Chinese economy and military might develop sufficiently to allow China to become globally preeminent. If this strategy is not derailed (and that is a major if), Swaine and Tellis argue, there are three plausible long-term scenarios: that China becomes domestically chaotic, internationally cooperative, or internationally assertive, perhaps to the point of global preeminence. Where the United States is the hegemon today and for the foreseeable future, China may seek to be the hegemon some decades hence.

Swaine and Tellis focus on hegemony based on economic and military power. However, as China may know from its earlier experience as a hegemon in East Asia, there is more to it than economic and military power. The United States became hegemonic in the West after World War II, when it possessed both economic and military power, and an attractive liberal ideology that provided profound economic and social benefits. It led a coalition against a militarily powerful, ideologically expansionist Soviet Union; when the Soviet Union collapsed, in part because its ideology was not functional in practice and its empire was held together by force, the United States found itself with a global stature of nearly hegemonic scope. Since 1991, U.S. ideology has reduced the opposition by other countries against the United States that might have been expected had the order it represents relied on military and economic power alone. While many voices, some in Mandarin, complain about U.S. hegemony, it provides a robust and
functional order in which states like China can develop in reasonable peace. Though it may not want to acknowledge the fact, China benefits from the hegemony of U.S. order, and from its enforcement by American naval and air power.

Unlike the United States after World War II, China does not appear to offer a globally compelling ideology; in the absence of such an ideology, increasing Chinese power and assertiveness may frighten countries to resist China by generating power internally or externally. Internal power might take the form of developing nuclear weapons. Most likely, external balancing would take the form of intensified ties with the United States, though Taiwan, Vietnam, and India all might have reasons to become better acquainted with each other; even a weak Russia might contribute to a China-constraining coalition. Increases in Chinese power—if only for defensive purposes, like looking after China’s energy interests abroad—might lead to resistance, because no country could be sure that Chinese intentions would not change. Counterbalancing would likely become more intense if China were to seek to move from regional hegemony to “geopolitical preeminence on a global scale.” The absence of an adequate consideration of such “balancing” may be the most significant weakness in this book. However, this is an impressive study of China’s grand strategy, and it is worthy of serious examination.

CARMEL DAVIS
University of Pennsylvania


In 1971 James Auer published The Postwar Rearmament of Japanese Maritime Forces, 1945–1971, “a book that [was] meant to challenge certain assumptions surrounding post–World War II Japan and its military, in particular its maritime forces.” Thirty years later, in post–Cold War Japan, some of those same suppositions persist. Of particular note, beliefs that Japanese rearmament is an American initiative, that Japan seeks only to provide for its own internal security while the United States is responsible for all external threats, and that militarism is returning to Japan, result in conflicts within Japan and among its Asian neighbors.

To many, the U.S.-Japan alliance both protects Japanese security interests and provides the cork that keeps Japanese militarism in the bottle. To students of Asia, the alliance follows a natural evolution resulting from the congruence of interests of two maritime nations as Japan reasserts its regional influence. To the prescient, the U.S.-Japan alliance is the bedrock for stability in a region rife with competing interests, developing conflicts, and unresolved animosities.

Peter Woolley, a professor of comparative politics at Fairleigh Dickinson University, has studied and written extensively on Japanese maritime development. He has written the next volume in this continuing story of Japan’s defense-system evolution, masterfully picking up where Auer left off. Woolley provides a concise analysis of the Japanese Defense Forces’ role in a world awakening to the transitional peace of the post–Cold War era.
His focus is an examination of the development and potential of the Japanese navy in the context of the U.S.-Japan Defense Security Alliance.

Rising from the ashes of World War II to become a legitimate regional power, Japan has developed the “second most powerful naval force in the world’s largest ocean.” This has led the United States and others to press Japan to contribute more to the security of the democratic and free-trading world, while its neighbors closely watch for signs of militarism that would signal resurgent nationalism and imperialism.

Through comparative analysis, Woolley presents the development of the Japan Maritime Self-Defense Force (JMSDF), and the Japan Self-Defense Forces (JSDF) generally, in decision-making models developed within each of the book’s seven chapters. He sees the efficacy of Japan’s military might and fear of its potency as catalysts that will awaken the long-dormant Japanese aim for regional hegemony and dominance.

Woolley begins by examining the cultural implications of the expanding role of the JMSDF, followed by the institutional dimension of the legal constraints imposed by the Japanese constitution on JSDF activities. In subsequent chapters, he traces the impact of strategic purpose on the development of these forces as Japan accepted the role of defending its sea-lanes out to one thousand miles. Woolley follows with an examination of the domestic and international implications of Japan’s expanding capabilities and operations as it sent minesweepers to the Persian Gulf in 1991, and how the incremental changes in Japan’s defense policy over the last three decades resulted in tacit international and domestic acceptance of its participation in peacekeeping operations. Woolley concludes with a valuable discussion of Japan’s changing defense posture and its relevance to aiding the United States in protecting shared vital interests in Asia.

Well researched and meticulously documented with an extensive bibliography, this book is an excellent reference for anyone wanting to understand Japanese defense policy and the forces that affect it. However, it is more than a book about the development of the JMSDF in the last three decades. It is also about the development of Japan’s national defense policy and the forces that move Japanese policy makers. It is a concise treatise that effectively uses maps and tables to help the reader understand key points.

Thus, the reader should be cautioned that the title of this book does not accurately convey its value as an examination of the strengths and weaknesses of the Japanese defense forces—not just the JMSDF—nor does it suggest the richness of the author’s analysis of Japanese policy making. This book provides much more, and it can serve as an excellent resource for gaining insight into the most important bilateral relationship of the United States.

ROBERT MORABITO
Commander, U.S. Navy
Naval War College


*Flying Black Ponies* is an effective combination of combat narrative, squadron history, and personal memoir, telling the story of Light Attack Squadron 4 (VAL 4, or the “Black Ponies”), a naval aviation squadron stationed in the Mekong Delta.
during the Vietnam War. During most of the war, the U.S. Navy made an intense effort with its Mobile Riverine Force to interdict enemy arms and supplies that flowed, primarily from Cambodia, across the Mekong Delta into the area surrounding Saigon. Kit Lavell’s book is a readable account of the Black Ponies’ important role in this hazardous interdiction campaign.

The Mekong Delta is not an easy place to conduct any type of military operation. It is a lush, steaming, tangled waterscape of swamps, soggy plains, and rice paddies crisscrossed with thousands of miles of rivers, streams, and canals. The canal-interdiction war in the Mekong was limited in 1968 in part by the inability to sustain close air support. The “Swift boats,” river patrol boats, SEAL teams, and the overstretched squadron of Navy attack helicopters (HAL 3, or the “Seawolves”) that were already in place clearly needed assistance. The new Commander, Naval Forces Vietnam, Vice Admiral Elmo R. Zumwalt, Jr., decided that air firepower was needed that could reach station quickly and remain there for several hours. The Black Ponies became operational in the Mekong during April 1969 in order to give fixed-wing aircraft punch, mainly with five-inch Zuni rockets, to the fight for control of the strategic river delta.

The book follows VAL 4 from its commissioning in San Diego, in January 1969, until its last mission and decommissioning in April 1972. The narrative also contains useful summaries of Navy tactics in the Mekong Delta, briefly describing MARKET TIME, GAME WARDEN, SEA LORDS, GIANT SLINGSHOT, DUFFLE BAG, ACTOV, and other operations in which the Black Ponies were involved. Lavell also weaves numerous other elements of the squadron’s experiences into his account. Lavell was one of the Black Pony pilots, flying 234 combat missions in the OV-10 Bronco aircraft during his tour with VAL 4. He effectively describes the frustration of being at the very end of the line of the Navy’s support and supply organization in Vietnam, and of dealing with rear-echelon staff administrators. He also pays deserved tribute to the enlisted maintenance and ordnance crews who worked long hours in the heat and humidity to keep the squadron aircraft safely airborne and armed. He portrays the sometimes humorous and somewhat disrespectful escapades of junior officers, particularly as the pullout from Vietnam neared in 1972.

Most importantly, however, this book is about the pilots of VAL 4 and their combat experiences. In three years of flying in the Mekong Delta, the squadron compiled a unique and impressive record. The Black Ponies’ ability to “scramble” when the Riverine Force made contact with the enemy continually resulted in direct air-ground firefights. The combat action is sharply drawn, evoking the tension, complexity, and confusion of delivering air strikes, especially in close proximity to friendly forces on the ground. When strikes took place at night and in bad weather (which they often did), the descriptions are even more harrowing. Several of the stories of the aerial fighting are effectively paired with personal memories of the same engagement by other personnel involved—river patrol boat crews, SEAL teams, and Riverine Force troops on the ground.

As Steven Coonts notes in his foreword, *Flying Black Ponies* can be read as a characteristic example of the way America fights its wars. Coonts describes the volunteers that fought the air war—young men of
blue-collar or decidedly middle-class background, mostly from farms or small towns. Their story is an important one, and the first-person accounts of individual sacrifice and aircrew heroism are a needed addition to the narrative of the Navy’s nearly forgotten war in the Mekong Delta.

WILLIAM M. CALHOUN
Naval War College


The issue of strategic bombing’s effectiveness is vitally important to political and military leaders. U.S. Air Force doctrine has argued for decades that airpower’s ability to operate directly and immediately at the strategic level of war is its unique and defining characteristic—a characteristic that must be exploited. Many disagree, so the debates have been long and heated.

Gian Gentile, a serving Army officer, now adds to the literature on this important subject. Unfortunately, he never really comes to grips with the key issue of effectiveness implied by the title of his book. Rather, he has chosen to repow some old ground, looking anew at the U.S. Strategic Bombing Survey (USSBS), chartered by President Franklin Roosevelt to examine and report on the effects of strategic bombing in World War II. Measuring bombing’s effectiveness and examining the workings of the USSBS that studied bomb effects are two different things.

The story of USSBS has been told before. In many ways it is a typical tale of wartime America. A need is identified, resources and personnel are mobilized, vast amounts of energy and material are expended, and notable gains are achieved. At the same time, the path to success is not a straight line—there is much inefficiency, debate, and compromise. Sausage is being made.

Gentile does not contest the findings of USSBS. Indeed, virtually no one has attempted to do so in the five-plus decades since they were released. The reason is simple: no one has the time, stamina, resources, or expertise to review the mountain of data collected and analyzed by the thousand individuals who conducted the USSBS.

Instead, Gentile seeks to discredit the survey’s findings by revealing flaws and inconsistencies in the survey itself. Primarily, he argues that Army Air Forces (AAF) leaders were so interested in forming a separate air force after the war that they induced bias into the USSBS. He is unable to make this charge stick. The survey was led by noted industrialists, bankers, economists, lawyers, and other professionals, most of whom had had little or no direct involvement with aviation prior to their work with the survey. Gentile admits that General Henry “Hap” Arnold, the AAF commander, stressed to these civilian leaders the need for objectivity, impartiality, and truth in the survey’s findings. Yet he treats such admonitions as duplicitous, despite the statement by the USSBS head that “at no time has there been the slightest inclination to interfere with us.”

In truth, it is difficult to imagine that men like John Kenneth Galbraith, Paul Nitze, and George Ball could have been manipulated and pressured to distort their findings. Common sense and logic tell us it is more likely that these men—and their hundreds of colleagues on the survey—examined thousands of
documents, interviewed hundreds of witnesses, visited scores of bombed sites, and then concluded that strategic bombing had indeed been a decisive factor in the Allied victory, as they reported. Alas, such a conclusion is unacceptable to Gentile. He must find nefarious schemes and schemers, and so he repeatedly questions the motives and veracity of the participants. For example, when General Curtis LeMay testified before Congress that he did not believe airpower could “win the war” and that a balanced mix of land, sea, and air forces was necessary for victory, Gentile dismisses his statement as a “shrewd and bureaucratically astute” tactic to manipulate his civilian superiors.

The USSBS has been controversial ever since it was written. Small wonder—attempting to measure the effects of strategic bombing in World War II was a massive undertaking, conducted at a time when the techniques of systems analysis were in their infancy. Gentile finds it troubling that survey members were not in total agreement. This should hardly come as a surprise. If the unfolding of historical events were simple and uncontested, our libraries would be far smaller.

His concluding chapter, dealing with the survey that analyzed the air campaign of the Persian Gulf War, is less tendentious. Here again, however, the author presents little that is new, and, more importantly, he does not attempt to address the book’s ostensible focus—the efficacy of strategic bombing.

Measuring the effectiveness of strategic air attack is one of the greatest challenges facing military planners today. It is an enormously complex and difficult problem that defies easy solution. Yet as airpower becomes increasingly dominant as a foreign-policy tool, such measurement is essential. This poorly reasoned and highly parochial book will not help us find answers to that pressing need, nor will it foster understanding among the services.

PHILLIP S. MEILINGER
Science Applications International Corporation
McLean, Va.


Lauris Norstad was a major Air Force leader during the defining years of the Cold War, and except for Dwight Eisenhower, he was the most prominent of all the Supreme Allied Commanders Europe (SACEUR) since that position was established in early 1951. Surprisingly, up to now, nothing definitive had been written on his role as SACEUR. Robert Jordan, a professor at the University of New Orleans and an authority on Nato, has filled that gap.

Norstad grew up in a small town in Minnesota and graduated from West Point in the class of 1930. Transferring to the Air Corps in 1931, he was one of that relatively small group of regular-officer aviators who provided air force leadership during World War II. When the war began, Major Norstad was serving on an air staff in Washington, D.C. He came to the attention of General Henry “Hap” Arnold, who headed what had become in June 1941 the Army Air Forces. In 1942 Arnold established a select group of young officers, the brightest he could find, to work in his immediate office. Norstad was one of them—he was on the way up.
That summer Arnold sent Norstad to England to serve as operations chief of the Twelfth Air Force—Eisenhower’s air arm for Operation TORCH, the November 1942 invasion of North Africa. In his war memoir, Eisenhower had this to say about Norstad: “[Norstad was] a young air officer who so impressed me by his alertness, grasp of problems, and personality that I never thereafter lost sight of him.” Before long, Brigadier General Norstad was operations chief for the Allied Air Forces Mediterranean. In the fall of 1944, Arnold returned Norstad to Washington, D.C., as chief of staff of the Twentieth Air Force, charged with planning the strategic bombing campaign against Japan. By war’s end, Major General Lauris Norstad, though not one of the top combat heroes of the Army Air Forces, was definitely one of its top comers.

From the end of World War II until the Korean War, the leadership of the reduced American armed forces struggled with new questions. Two of the most important were the role of nuclear weapons and how the U.S. military should be organized. With Eisenhower serving as chief of staff of the U.S. Army, and Norstad his deputy for operations (G-3), Norstad was involved in both issues, particularly in developing the compromises that led to the 1947 legislation resulting in the National Military Establishment, and ultimately to a separate Department of the Air Force. Subsequently, Lieutenant General Norstad, operations chief for the Air Force, played a major role in organizing the Berlin Airlift during the crisis of 1948–49. In the fall of 1950 he became commander of U.S. Air Forces Europe.

The main focus of Jordan’s book concerns Norstad’s subsequent twelve years of service in Europe, in particular the last six, when he served as SACEUR.

The author analyzes in detail three major issues confronted by Norstad that were most significant: the role and employment of nuclear weapons in alliance defense, the Berlin crises of 1958–59 and 1961–62, and the problem of balancing SACEUR’s roles as both an international and an American forces commander.

The nuclear weapons issue was complicated by the fact that the British had their own weapons, the French wanted theirs, and the West Germans, having none, were not quite certain they would be fully defended if the alliance had no recourse other than nuclear war. As Jordan demonstrates, Norstad was an able diplomat who succeeded in developing an alliance consensus on the role of nuclear weapons in deterring the Soviet Union.

Since Berlin was inside the Soviet-occupied zone of Germany, it became an ideal location for the Soviets to apply pressure on the alliance—by denying, or threatening to deny, access to the city. Though the issues were extremely complex, in a clear and interesting fashion Jordan explains Norstad’s role as diplomat and strategist in meeting Nikita Khrushchev’s challenge.

It is in his examination of SACEUR’s conflicting roles as an American and simultaneously an international commander that Jordan makes his major contribution. This issue came to a head for Norstad with the arrival of the Kennedy administration in 1961, with its secretary of defense, Robert Strange McNamara. The substantive issue was the nature, role, and control of nuclear weapons as an element of Nato strategy. The process issue was that the administration found it hard to accept SACEUR’s dual role, tending to
view Norstad as an American commander only. The details cannot be developed within the confines of a review, but in the end Norstad was forced to walk the plank—though the final jump was delayed for a period of two months by the administration’s need for his assistance during the 1962 Cuban missile crisis.

Robert Jordan has produced an important work that is thoroughly researched, nicely written, and most insightful. No doubt it will be the definitive biography of Lauris Norstad—Cold War airman, strategist, and diplomat. The book will also be of interest to those involved in the study of civil-military relations, especially in these years of increased commitment of U.S. military forces in multinational or international interventions.

DOUGLAS KINNARD
Brigadier General, U.S. Army, Retired
Emeritus Professor of Political Science
University of Vermont


Conrad Crane is a research professor for military strategy at the Strategic Studies Institute, U.S. Army War College, and formerly a professor of military history at the U.S. Military Academy. Crane previously wrote Bombs, Cities, and Civilians: American Airpower Strategy in World War II (1993), which is widely respected for its rich and adroit analysis. American Airpower Strategy in Korea, 1950–1953 is a comprehensive, thoroughly researched treatment of the many issues that the newly constituted U.S. Air Force faced as a result of having to fight its first war as an independent service—a war that it was not doctrinally or materially prepared for, and that the service had neither anticipated nor especially wanted to fight. Crane logically takes the reader through the war from the prehostilities period, which generally set the stage for the limited character of the war and specifically established the character of the Air Force’s contribution; the opening moves and initial setbacks; the miraculous end-around at Inchon and subsequent march to the Yalu; the bitter winter of 1950–51; and finally to the stalemate along the thirty-eighth parallel.

Crane analyzes the performance of the Air Force in conducting air warfare in a regional, limited conflict at a time when the service was focused on strategic nuclear war and restricted by government policy as to the resources that could be allocated to Korea. It was a condition that the Air Force would again confront in Vietnam. The Korean War presented the Air Force with a myriad of challenges, not the least of which was the attempt to meet high expectations for operational effectiveness based on results obtained during World War II.

However, the very nature of the new conflict constrained that effectiveness. A classic example of the limited nature of the Korean War was the prohibition against crossing the Yalu River to engage enemy forces or interdict lines of communication. Crane also takes great pains to highlight how austere were the resources made available to the Korean area of operations, because the Air Force was required to maintain the bulk of the active component in a ready status to respond to other worldwide threats. This requirement was the catalyst for many issues that arose during the conduct of the war, among them the decision to recall to active duty large numbers of aircrewmen who had served in World War II and
were in many cases not keen to leave their families and jobs to serve in an undeclared “police action.”

In addition, Crane recognizes, the Air Force was challenged by interservice rivalry with the Army and the misunderstanding of its role in battlefield air interdiction, and to a lesser degree by negative perceptions created by strategic bombing at the expense of close air support for the Army. He points out that the frustration felt by Army commanders was exacerbated by the effective and dedicated close-air support provided to the Marines by their air component. The Army continually questioned why the Air Force could not provide for it the same level of effective support.

Crane also rightly recognizes the effective leadership and operational genius of the Far East Air Forces (FEAF) planner, Brigadier General Jacob Smart, who was able to produce a coherent interdiction strategy which he skillfully “sold” to the Army. Smart recognized the difficulty of conducting tactical interdiction operations against an entrenched enemy who did not require much in the way of supply. He reoriented FEAF’s interdiction efforts away from cutting tactical lines of communications to striking such operational targets as hydroelectric facilities, supply distribution centers, and other “deep” targets, all with some effect.

Crane’s book is a valuable compilation of the contributions of the Air Force in the prosecution of the Korean War. Crane reveals the warts but also gives glowing credit where it is due. Much more than a mere chronology, this is an insightful book that is a must-read for critical students of this conflict.

WILFRED F. BROWN
Colonel, U.S. Army
Naval War College


For thirty years after the Second World War, historians said very little about the role of signals intelligence in the Battle of the Atlantic—because either they did not know about the Allies’ remarkable code-breaking successes, or they could not write about what they did know. That changed in 1974, when revelations about ULTRA exposed the full extent of the Allied penetration of Germany’s signals. Unfortunately, in the subsequent rush to rewrite the history books to include ULTRA, its significance was frequently inflated.

In this study of ULTRA and the Battle of the Atlantic, Gardner offers the most fully developed case yet that monocausal explanations for the Allied victory in this campaign are inadequate—that ULTRA was only one critical factor among many. Gardner provides two case studies to support his argument. The first demonstrates that Britain’s growing ascendancy over the U-boat in 1941 had many causes, most of which were unconnected with ULTRA. Just as important as code breaking, if not more so, was the tightening up of the convoy system and the German decision to shift U-boat operations westward in order to avoid the increasingly hostile environment around the British Isles. The greater availability of escort vessels and growing American assistance also played an important part in turning the tide in Britain’s favor in 1941. It is therefore a mistake to suggest, as some have done, that ULTRA alone may have saved as much as two million tons of shipping during this period.
Gardner’s second case study covers the period from mid-1942 to mid-1943, the final turning point in the campaign. Once again, he makes a strong argument that factors other than ULTRA were essential for the Allies’ success, most importantly the closing of the mid-Atlantic “air gap” and the increasing number and effectiveness of Allied escort vessels. By 1943, Gardner notes, it was more hazardous than ever for the Germans to attack convoys. At the same time, the growing number of U-boats operating in the mid-Atlantic made the evasive routing of convoys extremely difficult for the Allies, even when ULTRA was available on a timely basis.

These case studies cover periods when ULTRA was most valuable because Germany was employing its U-boats in “wolf pack” attacks on convoys. When U-boats operated individually rather than in groups, which was the case for nearly two-thirds of the war, ULTRA’s value was much less. Gardner also emphasizes that there were lengthy periods when German signals were being read only after considerable delay. For much of the war, he concludes, ULTRA’s principal use was to enable the Allies to build up a general picture of the size and methods of the German U-boat fleet. It seldom had any impact on the Battle of the Atlantic at the operational or tactical levels.

This book deliberately focuses on the turning points of the campaign and the broad relationship of ULTRA to other factors; it does not provide either a blow-by-blow account of the Battle of the Atlantic or a systematic examination of ULTRA’s employment by Allied commanders. The period from June 1943 until the end of the war is largely ignored. However, Gardner devotes a considerable portion of this book to background information about the Battle of the Atlantic, including the importance of economic factors, the role of technology, Germany’s own code-breaking activities, and the workings of convoy. This material should be useful for the general reader even if it offers little for the specialist.

Gardner’s broad conclusions are carefully reasoned and well balanced. The Battle of the Atlantic would have been a harder and costlier struggle without ULTRA, but the Allied victory cannot be solely attributed to code-breaking successes. Decoding History will not be the final word on signals intelligence and the Battle of the Atlantic, but it may put an end to the wildly exaggerated claims that are sometimes made for ULTRA.

CHRISTOPHER BELL
Naval War College


There are few untold stories left from World War II, but the actions of the Royal Navy’s Coastal Forces can be described as little known and unappreciated. Serving in small plywood craft much like the more famous American PT boats, the Coastal Forces wreaked havoc with Axis forces in British and foreign waters. Operating from small harbors and primitive forward bases, theirs was a war of small, close-knit crews and close action with the enemy. They did it all, from convoy escort to shipping interdiction, clandestine landings to reconnaissance operations, and finally, distant screening for invasion forces. Coastal Forces were a ubiquitous presence in the
European theater. Wherever there were Axis forces in coastal waters or areas, the Royal Navy dispatched Coastal Forces to counter them. Yet strangely, little has been published about these deadly fast-attack craft and their courageous crews. Dog Boats at War redresses some of that shortfall in naval literature.

Written by a wartime motor gunboat commander, Leonard C. Reynolds, this book tells the Coastal Forces’ story with an authenticity that can only be produced by one who was there. Reynolds focuses on the actions that took place during his service in 1942–45, and on the class of boat in which he served, the Fairmile D-class motor-torpedo and motor gunboat (MTB and MGB, respectively). The title, Dog Boats at War, is derived from the nickname given to the rather blunt-looking Fairmile fast-attack craft.

The Royal Navy entered World War II with two flotillas of underpowered MTBs. They proved woefully inadequate against the German Schnellbooten and Italian MAS torpedo boats. The Admiralty tasked Noel Macklin of Fairmile to develop a fast-attack boat to compete with the enemy boats. Macklin’s design was ready by March 1941 and was put into production six months later after a rushed but successful testing program. The first boats entered service in the English Channel by early spring 1942. Originally intended as motor gunboats for convoy escort, they were converted to torpedo boats while under construction. Equipped with two twenty-one-inch torpedo tubes instead of a six-pounder cannon, the MTBs proved very effective at intercepting German convoys transiting off the Dutch and French coasts.

The “dog boat” was a simple and robust design, but its performance was not extraordinary. Its 115-foot hull had a blunt semi-hard-chine design and used four Packard 1,250-horsepower engines to achieve a maximum speed of thirty-two knots—some five knots slower than its German or American counterparts. The dog boats were also small in comparison to their opponents but rode better in a seaway. Their superior stability often proved decisive in the rough waters of the English Channel and off the Norwegian coast. In the end, however, it was the crews that made the difference. It took a special kind of sailor to man a plywood boat filled with five thousand gallons of highly volatile aviation gasoline. The Royal Navy found their early crews among the yachtsmen, racing enthusiasts, and fishermen who populated Britain’s coastal towns and villages. These early recruits shaped the force that followed. The dog boats went to war with small crews of thirty to thirty-two men commanded by young Royal Navy Volunteer Reserve officers, many drawn from the Commonwealth. The commanding officers were generally in their twenties, and most had at least two years of experience in fast-attack craft. The crews were also either reservists or “wartime only” ratings. The average age of the crewmen was between eighteen and nineteen, with perhaps one career regular petty officer on board to provide experience. The result was a tightly knit, young, and aggressive crew—ideally suited to the unconventional fighting that characterized warfare in fast-attack craft. Survival depended on the commander’s judgment and the crew’s teamwork. It was a successful formula that served the boats well from Norway to the eastern Mediterranean. The Royal Navy’s Coastal Forces operated in a style not unlike that of light
forces of the Napoleonic era. Some of the actions described here are worthy of the best adventure fiction. Many boats ran agents, supplies, and weapons to the underground forces resisting Nazi occupation. For example, the 30th MTB Flotilla’s boats often hid in caves along the Norwegian fjords or crept along dark “leads,” dodging German patrols while seeking contacts and recruits among the occupied population. MTBs carried the kidnapped German general Werner Kreipe off Crete and transported him to Egypt for interrogation. They also played a deadly game of hide and seek with the Axis navies and the Luftwaffe in the Aegean Sea and among Yugoslavia’s coastal islands. Wherever they served, the dog boats were the force of choice for engaging the enemy closely—and they paid dearly for it, losing 273 officers and men killed in action.

Some 228 dog boats were built between November 1941 and April 1945. They fought in over three hundred actions, sinking and damaging innumerable Axis vessels while losing some thirty-seven of their own. On the basis of eight years of research in official records and interviewing people involved, Reynolds has compiled as complete and accurate a record of the dog boats’ actions as humanly possible. *Dog Boats at War* is a brilliant, if occasionally dry, treatment of an important and all but ignored part of the Royal Navy’s history in World War II. It is worth its price, and I hope it will be followed by similar works on the Coastal Forces’ other elements.

**Carl Otis Schuster**  
*Captain, U.S. Navy, Retired*  
*Sailua, Hawaii*  

---


The Navy has had a remarkable and productive group of in-house research and development laboratories. Sadly, with the drawdowns of the post–Cold War era, many have been closed, among them the “White Oak Laboratory”—the Naval Ordnance Laboratory at White Oak, Maryland. WOL, as it was popularly known, gave the Navy a tremendous legacy of technology, weapons, and people. Fortunately, that legacy has been preserved by William Anspacher, Betty Gay, Donald Marlowe, Paul Morgan, and Samuel Raff in this richly detailed account of the laboratory’s history.

First, the required disclaimer: this reviewer spent twenty years of his midcareer with WOL. And good years they were. The laboratory was built in 1946 in what was then remote suburban Maryland, where ordnance testing would presumably not disturb the neighbors. It was an outgrowth of the Mine Building at the Washington Navy Yard, and mine development was the core of its original work. From it came the Navy’s postwar mines: the Mark 50 series, CAPTOR, the submarine-launched mobile mine, and the Destructor series.

In 1948, high-mach-number wind tunnels captured at Peenemunde, Germany, were installed at WOL and became operational. With these, the laboratory began a new line of technological development for the Navy. From those first tunnels grew a series of hypersonic wind tunnels that gave engineers the ability to test re-entry vehicles at speeds up to Mach 14.
In addition to military work, the tunnels were used by the space shuttle program to study reentry forces. The laboratory became the nation’s center of excellence in hypersonic aerodynamics.

White Oak was never a “big systems” center, and in that lay the roots of its ultimate demise. Rather, it was a technology center, focusing on ordnance and system components.

From 1946 on, the laboratory created extensive expertise for the Navy in explosives, warhead design, fuzing, metallic and nonmetallic materials, magnetic silencing, nuclear weapons effects, and underwater acoustics. The authors devote substantial chapters to each of these—the people, the anecdotes, the products, and the fleet applications. Describing them all is beyond the scope of this review; that pleasure is saved for the reader. Fleet-savvy readers will recognize many WOL-developed components in the systems they use today.

Magnetic-silencing research necessitated the construction of a unique building made entirely of wood and nonmagnetic metals. Up close it looked more like the work of a cabinetmaker than a government-contracted building. Magnetic-signature and degaussing work at WOL led to the development and fielding of the drive-through deperming facilities for submarines at Kings Bay, Georgia, and Bangor, Washington.

The Naval Science Assistance Program, under whose aegis laboratory scientists were assigned to major fleet commands to solve technical problems and introduce new technical concepts, was created and managed at WOL. Two generations of Navy scientists went to sea, wrung salt water out of their socks, gave the Navy new tricks, and returned with solid understandings of their ultimate customer.

Technology developed at WOL spun off new and unanticipated applications in the civilian world. Nitinol, a metal alloy with temperature-stimulated memory properties, found use in orthodontics. Research in the mathematics of nonlinear systems led to techniques for controlling heart arrhythmia, making a chaotic heartbeat a regular one.

The legacy of the laboratory for the fleet is in the technology and hardware now deployed; the legacy in the hearts of the alumni and alumnae (and there are many of the latter) is in the people and the images. Fortunately, the authors have done them all full service. The book is generously illustrated and filled with the people, from the recipient of two Nobel Prizes in physics to the fellow who liberated a fire truck from another government installation—White Oak didn’t have one.

Beyond being a fine institutional history, the book is a valuable study in public administration as practiced for military research and development. The authors have unearthed and analyzed an impressive amount of bureaucratic history involving all the players in the Navy’s research and development hierarchy.

In their analysis, the White Oak Laboratory suffered from not having a platform-based mission. As an ordnance and technology laboratory, it was not the creature of any of the Navy’s controlling baronies—air, surface, or submarine. The laboratory served them all but had the prime responsibility for no platform-based major weapon system. Thus it was always just outside the door, looking through the window but not sitting at the table.
In 1974 the White Oak Laboratory was merged with the Naval Weapons Laboratory at Dahlgren, Virginia, to become half of the Naval Surface Warfare Center, and it found itself in the surface warfare community. A long, painful, but inexorable decline in the laboratory’s fortunes began. The authors’ detailed and insightful treatment of this period, with all its bureaucratic infighting and personalities, is an important part of the book. At its core, the lesson in public administration is that pure technology and elegant components alone are not a sufficient raison d’être in the military research and development world. Such an institution must have a clearly defined customer base and serve it with comprehensive integrated systems.

In 1996, the White Oak Laboratory was closed, and its people and projects were sent to other naval centers. At this writing the grounds and main buildings are expected to become new laboratories for the U.S. Food and Drug Administration. This is not unfitting. A plaque in the lobby will remind everyone of what was once accomplished there for the Navy.

FRANK C. MAHNCKE
Washington, D.C.
FROM THE EDITORS

In Memoriam

MICHAEL I. HANDEL
The editors greatly regret to report the passing on 14 June 2001 of an esteemed colleague, scholar, friend, and contributor to this journal—Dr. Michael I. Handel, a member of the Naval War College’s Department of Strategy and Policy and the holder of the Philip A. Crowl Chair of Comparative Strategy.

ANDREW E. GIBSON
The editors further regret to report the passing on 8 July 2001 of Andrew E. Gibson, formerly the holder of the Emory S. Land Chair of Maritime Affairs at the Naval War College, Assistant Secretary of Commerce, president of the Delta Steamship Company, coauthor of The Abandoned Ocean, and in World War II the youngest master (age twenty-two) of a Liberty ship, commanding the Leonidas Merritt. He also was a frequent contributor to the Naval War College Review.

NEWPORT PAPER NUMBER 15
A new volume in our Newport Papers series is now available: International Environmental Law and Naval War: The Effect of Marine Safety and Pollution Conventions during International Armed Conflict, by Sonja Ann Jozef Boelaert-Suominen. Dr. Boelaert-Suominen is currently an adviser on international law to the Office of the Prosecutor of the International Criminal Tribunal for the Former Yugoslavia. For a copy, contact our associate editor, Mrs. Pat Goodrich, at goodricp@nwc.navy.mil, (401) 841-6583, or by fax at (401) 841-1071. The monograph is also available at full length (Adobe Acrobat required) on our Website: www.nwc.navy.mil/press/npapers/newpaper.htm.
WINNERS OF THE HUGH G. NOTT PRIZE FOR 2000
The President of the Naval War College has announced the winners of prizes for the finest articles (less those considered for our history prize) appearing in the Naval War College Review in the 2000 publishing year:

First Prize ($500)  Dr. John Garofano, “Deciding on Military Intervention: What Is the Role of Senior Military Leaders?” (Spring)
Second Prize ($300)  Yoram Dinstein, “The Right to Humanitarian Assistance” (Autumn)

WINNER OF THE EDWARD S. MILLER HISTORY PRIZE FOR 2000
Through the generosity of the distinguished historian Edward S. Miller, the President of the Naval War College has awarded a prize to the author of the finest article on a historical subject to appear in the Naval War College Review in 2000. The winner ($500) is Dr. Donald Chisholm, of the Naval War College, for “Negotiated Joint Command Relationships: Korean War Amphibious Operations, 1950” (Spring).
OF SPECIAL INTEREST

FELLOWSHIPS AND GRANTS AVAILABLE
In order to encourage the study of American naval history, the Naval Historical Center offers research support for established scholars and doctoral candidates. For 2002–2003, the following opportunities are available (for information and forms, consult the NHC Website at www.history.mil):

Rear Admiral John D. Hayes Predoctoral Fellowship: $10,000 to a predoctoral candidate who is undertaking research and writing on a dissertation in the field of U.S. naval history. Applicants must be U.S. citizens enrolled in an accredited graduate school who will have completed all requirements for the Ph.D. except the dissertation by 1 June 2002. Deadline for applications is 28 February 2002.

Vice Admiral Edwin B. Hooper Research Grants: Up to $2,500 each to individuals undertaking research and writing in the field of U.S. naval history. Applicants should have either Ph.D.s or equivalent credentials, and they must be U.S. citizens. Deadline for applications is 28 February 2002.

CALL FOR PAPERS
The Society for Military History will hold its sixty-ninth annual conference at the Monona Terrace, Madison, Wisconsin, 4–7 April 2002, on the theme “War and Remembrance: Constructing the Military Past and Future.” The program committee particularly invites proposals for papers and panels that assess the military classics, memoirs and reminiscences, military reformers, and military leadership. Proposals for papers and panels treating all aspects of military history are welcome, as always. Proposals should include a one-page abstract for each paper, outlining topic, thesis, and sources, and should enclose a brief c.v. The program committee intends to post the abstracts on the SMH Web site, www.smh-hq.org. The committee welcomes volunteers to serve as chairs or commentators. Please submit proposals for papers and full panels no later than 1 November 2001, to: Prof. Jerry Cooper, Dept. of History, University of Missouri–St. Louis, St. Louis, Mo., 63121; tel. (314) 516-5735, fax (314) 516-5781, e-mail cooper@msc.umsl.edu.