NATO's Future Conventional Defense Strategy in Central Europe

Theater Employment Doctrine for the Post–Cold War Era

Richard L. Kugler

DISTRIBUTION STATEMENT A:
Approved for Public Release -
Distribution Unlimited

20041208 064

RAND | ARROYO CENTER

The research described in this report was sponsored by the United States Army, Contract No. MDA903-91-C-0006.

Library of Congress Cataloging in Publication Data

Kugler, Richard L.

NATO's future conventional defense strategy in Central Europe : theater employment doctrine for the post-cold war era / Richard L. Kugler.

p. cm.

"Prepared for the United States Army."

Includes bibliographical references.

"R-4084-A."

ISBN 0-8330-1188-X

1. North Atlantic Treaty Organization—Armed Forces. 2. Military doctrine—Europe. 3. Military doctrine—North America. 4. Central Europe—Defenses. I. RAND Corporation. II. United States. Army. III. Title.

UA646.3.K85 1992

355'.031'091821—dc20 91-32786

CIP

The RAND Publication Series: The Report is the principal publication documenting and transmitting RAND's major research findings and final research results. The RAND Note reports other outputs of sponsored research for general distribution. Publications of RAND do not necessarily reflect the opinions or policies of the sponsors of RAND research.

Published 1992 by RAND
1700 Main Street, P.O. Box 2138, Santa Monica, CA 90407-2138

R-4084-A

NATO's Future Conventional Defense Strategy in Central Europe

Theater Employment Doctrine for the Post–Cold War Era

Richard L. Kugler

Prepared for the
United States Army

RAND

Approved for public release; distribution unlimited

PREFACE

This study was written within the Arroyo Center's Policy and Strategy Program, as part of a project sponsored by TRADOC on the "Future of the U.S. Army in Europe." It provides a political-military analysis for thinking about how NATO's conventional defense strategy can be adjusted to contribute to stability in Central Europe in the coming post–Cold War era. In particular, it focuses on the new mission of defending the borders of a unified Germany after Soviet forces have entirely withdrawn from Eastern Europe. Much of the analysis was made available to the Army and the Defense Department in early 1990, and it was updated in February 1991. The study thus was completed before the Soviet Union was disestablished in late 1991. Although the specific threat postulated by this study has been invalidated by the Soviet Union's demise, the tasks of defending NATO's borders and maintaining a military balance of power in Europe remain on the alliance's agenda. For these reasons, the employment issues raised and the conceptual framework developed should remain relevant for some time. This study should be of interest to U.S. government officials responsible for NATO policy and to anyone concerned about the future of the European security system.

THE ARROYO CENTER

The Arroyo Center is the U.S. Army's federally funded research and development center (FFRDC) for studies and analysis operated by RAND. The Arroyo Center provides the Army with objective, independent analytic research on major policy and organizational concerns, emphasizing mid- and long-term problems. Its research is carried out in four programs: Strategy and Doctrine; Force Development and Technology; Military Logistics; and Manpower and Training.

Army Regulation 5-21 contains basic policy for the conduct of the Arroyo Center. The Army provides continuing guidance and oversight through the Arroyo Center Policy Committee (ACPC), which is co-chaired by the Vice Chief of Staff and by the Assistant Secretary for Research, Development, and Acquisition. Arroyo Center work is performed under contract MDA903-91-C-0006.

The Arroyo Center is housed in RAND's Army Research Division. RAND is a private, nonprofit institution that conducts analytic research on a wide range of public policy matters affecting the nation's security and welfare.

Lynn E. Davis is Vice President for the Army Research Division and Director of the Arroyo Center. Those interested in further information about the Arroyo Center should contact her office directly:

Lynn E. Davis
RAND
1700 Main Street
P.O. Box 2138
Santa Monica, CA 90407-2138

Telephone: (310) 393-0411

SUMMARY

The London Summit of July 1990 marked a watershed for the Western alliance. NATO's political leaders acknowledged that the Cold War's end should profoundly alter the way NATO thinks about security and defense, and they decided to prepare a new military strategy to replace flexible response (MC 14/3). In taking this major step, they launched NATO on the uncertain but necessary path of redefining its military raison d'être for the coming era.[1]

This positive feat notwithstanding, the London Summit also left a strategy conundrum in its wake. In declaring that nuclear forces are to become weapons of "last resort," it indirectly called on NATO to maintain a conventional defense that by itself can blunt nonnuclear aggression. It decreed that NATO's forces should remain mobile, versatile, and capable of responding flexibly to the crises of the future. But it also mandated that NATO's force posture is to be smaller, less ready, and more reliant on mobilization than during the Cold War. NATO's conventional defenses are to assume a larger role in its future military strategy, but with a posture that is less well-primed than before.

This study examines NATO's future conventional defense prospects in Central Europe and addresses the issue of how this puzzle can be solved. It offers no fixed blueprint; it provides an analytical framework for thinking about the challenges, problems, alternatives, and tradeoffs ahead. It analyzes NATO's future "theater employment doctrine": the way NATO uses military forces on the battlefield to attain its goals. Employment doctrine provides the critical link between strategy and force posture. This study endeavors to shed light on how NATO can attain the demanding goals of its new strategy in the upcoming era of lower military preparedness.[2]

A fundamental premise of this study is that NATO will need to remain militarily prepared to preserve its security and respond to future challenges. Another premise is that NATO's defense planning will need to address possible contingencies involving Soviet aggression against its borders. To be sure, this is only one of the many and diverse contingencies that could confront NATO, but it has important implications for NATO's defense strategy, forces, and command relationships. It also promises to have a major effect on NATO's employment doctrine.

This study does not postulate, however, that the future in Europe will produce a tense and confrontational relationship with the USSR. The London Summit Declaration and the Charter of Paris both proclaim that the Soviet Union is no longer an adversary of the West and that NATO's nations aspire to normal diplomatic relations with it. Recognizing these goals and the West's hope for a peaceful Europe, this study conducts what amounts to an exercise in prudent defense planning. It addresses the kinds of military plans that NATO will need to maintain as an insurance policy, even as NATO's nations endeavor to build a truly cooperative order in Europe.

This study addresses the following question: What are the options for NATO's theater employment doctrine of the future, and how should they be appraised in political-military

[1]See London Summit "Final Declaration," reprinted in The *New York Times*, "Text of the Declaration after the NATO Talks," July 7, 1990, p. 5.

[2]For a treatment of NATO's future role in Europe that covers security issues beyond military strategy, see Kugler, 1990b.

terms? During the Cold War, NATO's doctrine has been one of linear defense at the old inter-German border.[3] If linear defense no longer is feasible because NATO's forces will be smaller, then what kind of doctrine should NATO adopt? And if NATO no longer will be defending on the inter-German border because Germany is now unified, then where should it defend? What do the answers to these questions suggest about NATO's force levels and buildup rates and about the equally thorny issue of how coalition defense is to be conducted?

NATO will need to adopt a new, quite different theater employment doctrine, one that defends further eastward and more flexibly than linear defense contemplated. The task of designing a new doctrine will not be easy, and will require NATO to remain vigilant about future defense requirements. All of the viable alternatives will require a major uprooting of NATO's long-established defense practices. NATO will need to learn how to mobilize differently and to fight differently. Changes will have to be made not only in NATO's force posture, but also in how coalition defense is conducted. In particular, NATO's "layer cake" array will have to give way to something new. Nevertheless, these changes can be orderly and well-managed, leaving NATO with a viable conventional defense strategy even if forces are somewhat smaller than they are now. The key is a concerted planning effort that forges a coherent relationship among NATO's future defense strategy, employment doctrine, and force posture.

CRITICAL FACTORS AFFECTING NATO'S PLANNING

Receding of the Soviet/Warsaw Pact Military Threat

The impending withdrawal of Soviet forces from Eastern Europe and the collapse of the Warsaw Pact will sharply reduce the military threat facing NATO in Central Europe. In the years ahead, NATO no longer will have to worry about a massive (90-division) attack on its territory launched after less than one month of enemy mobilization. A smaller threat and a longer period of mobilization will ease NATO's traditional defense planning dilemmas.

Nevertheless, NATO will continue to face a serious defense agenda in Central Europe (as well as on the flanks). The future European security system most likely will have unstable fault lines of its own. Barring an internal collapse that renders it unable to function as a cohesive nation-state, the Soviet Union will probably remain a military superpower to be reckoned with. The USSR's goals, while less ideological and expansionist than before, still could come into conflict with those of the West. If so, NATO's future security policy will probably be cast in terms of maintaining a stable military balance of power and security architecture in Central Europe. This policy will require NATO to maintain a military counterweight to the USSR, one that could defend against a Soviet force of about 50–65 divisions that could be mobilized and deployed into Central Europe within some 60–90 days. This "threat" is smaller and slower building than before, but it is not a matter to be discounted in NATO's planning.

Effect of Force Reductions on Linear Defense

Within five years, NATO's forces themselves are likely to be reduced by 10–25 percent, perhaps more. These reductions will undercut NATO's ability to continue executing its

[3]"Linear defense" is defined as an employment doctrine that commits most of NATO's ground combat forces to forming a strong frontal wall directly on NATO's borders, with few forces withheld as theater reserves.

traditional linear deployment doctrine along either the old inter-German border or Germany's equally long new border. A linear doctrine would draw so many forces into the frontal zone that NATO's posture would lack the reserves that often are needed to contain an enemy breakthrough attack. As a result, NATO could still be vulnerable to conventional defeat even if the future conventional balance is one of parity, or something like it.

To offset this problem, NATO would need to craft a less rigid and brittle employment doctrine. A variety of options are available, including the idea of a "nonlinear" defense that would shift NATO's center of gravity further to the rear. All of these options carry tradeoffs, and each would impose its own new, unique demands on NATO's defense posture and plans. The task ahead is to evaluate these options carefully and then to forge a synthesis of them that makes sense for NATO at each stage of the evolving European defense situation.

Germany's Unification

The North Atlantic Treaty commits the Alliance to defending the territory of its members. Assuming the German government so desires, the unification of Germany will compel NATO to extend defense coverage over eastern Germany, the former German Democratic Republic (GDR). The withdrawal of Soviet forces from German territory by 1995 will permit NATO to contemplate extending its security umbrella in this way, but the German-Soviet Treaty of late 1990 will prevent NATO from physically deploying its forces into eastern Germany in peacetime. NATO must therefore craft a new theater employment doctrine calling for its forces to deploy into eastern Germany in a crisis and to take up positions there.

This change promises to introduce an entirely new dimension into NATO's defense plans for crisis management and wartime defense. After 1995, the military balance in Central Europe will take the form of two armies, less ready than before, facing each other across a lengthy buffer zone, with Poland in the middle. NATO's ability to maintain stability in this situation will depend upon its success at designing a force posture, employment doctrine, and mobilization capacity that enable NATO to defend not only traditional Cold War borders, but eastern Germany as well.

NATO'S Political-Military Goals

NATO's future employment doctrine clearly will depend upon whether the Alliance, in fact, decides to strive for a military posture that meets traditional standards for deterrence and defense. Beyond this, NATO's doctrine will be affected by how it decides to preserve its own internal unity. NATO's commitment to a viable conventional defense posture in Central Europe will be necessary to ensure that Germany remains satisfied with its membership in the Alliance. It also will be needed to ensure that the United States remains comfortable with the idea of providing extended nuclear deterrence, and that NATO's other members see safety in a military strategy that makes nuclear weapons a last resort.

These political imperatives do not mean retention of NATO's traditional layer cake array, which committed the forces of five different nations to front-line roles. Indeed, the military and political dynamics of the new European defense situation argue for a different approach to fostering coalition planning. Figure S.1 illustrates how the layer cake would appear if it were transposed onto Germany's new eastern border. NATO would have few reserves, British and Belgian forces would protect Berlin, and NATO's most powerful German and U.S. forces would be deployed on terrain that is unlikely to be attacked. Once

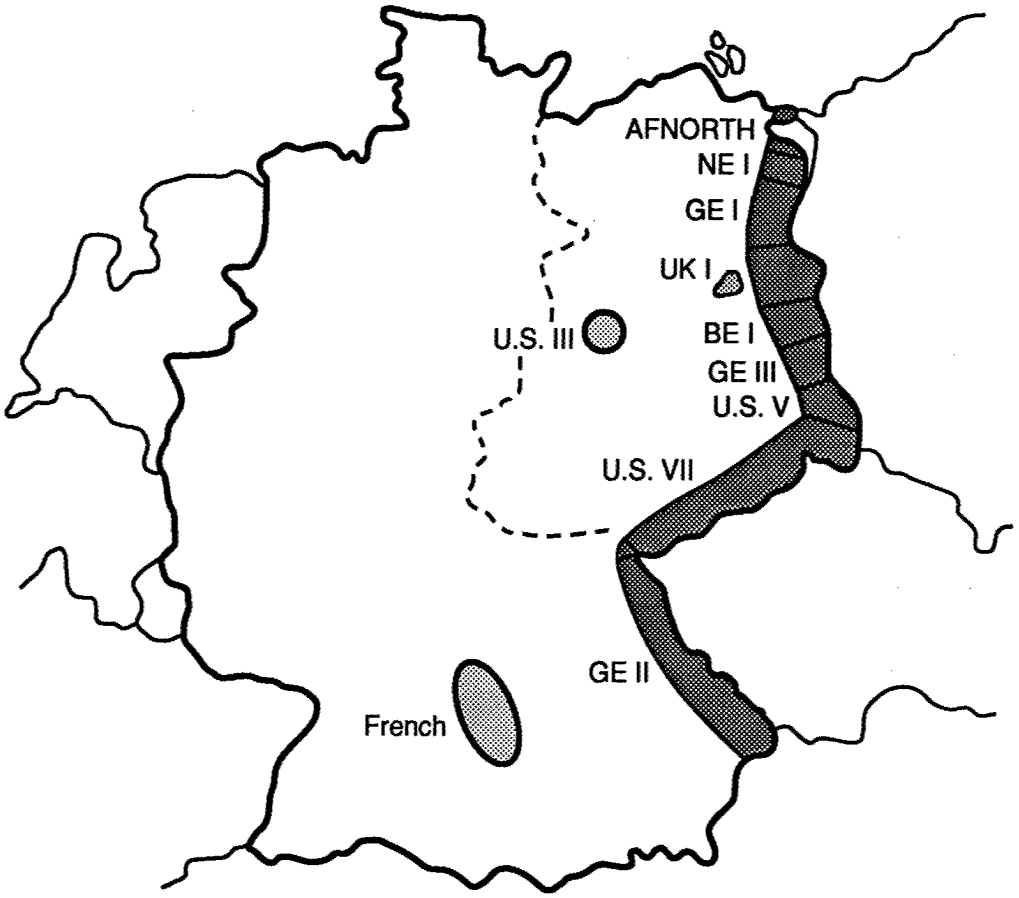

Fig. S.1—NATO's Layer Cake Array No Longer Will Be Appropriate

NATO's task becomes that of defending eastern Germany, a new scheme for distributing national responsibilities will need to be designed, one that makes military sense but still safeguards NATO's unity.

THEATER EMPLOYMENT DOCTRINES FOR THE FUTURE

These considerations suggest that NATO should alter its future employment doctrine in ways that conform to the three stages of Central Europe's evolution:

- The near term, during which large Soviet forces will remain in eastern Germany, thereby conceivably permitting a reinforced attack across the old inter-German border.
- The transitional mid-term (1993–1994), during which some Soviet forces will remain, but NATO will be able to contemplate defense at least of some eastern German territory.

- The "end game" (1995 and later), after Soviet forces have entirely withdrawn and defense of all of eastern Germany will become possible.

Near Term

Figure S.2 displays how NATO's employment doctrine might be altered in the near term to make its present linear defense at the old inter-German border less vulnerable. There would be a still-strong emphasis on forward defense, but in a more flexible manner. The employment doctrine envisioned here would shift NATO's center of gravity northward to protect the exposed North German Plain by sending CENTAG reserves to NORTHAG in a crisis. NATO would ease CENTAG's defense task by extending its front line to cover the densely forested Thuringerwald. Meanwhile, NORTHAG's front line forces would defend more flexibly, either by trading space for time or by adopting nonlinear tactics that call for mobile operations from their corps rear areas.

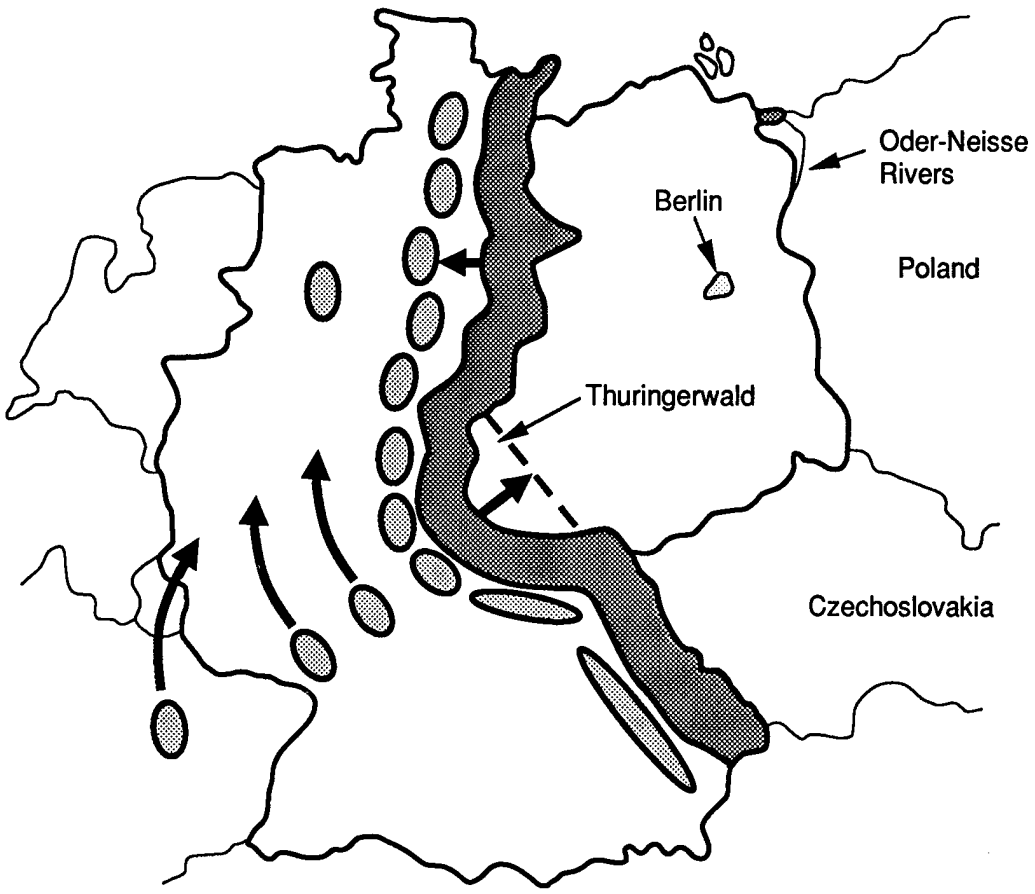

Fig. S.2—Near Term: A More Flexible Defense of the Old Inter-German Border

Mid-Term

In the mid-term, as Fig. S.3 suggests, NATO's employment doctrine would begin to move eastward. In a crisis, Soviet forces still based in eastern Germany would probably retreat into an enclave along the Oder-Neisse rivers, where they would await reinforcement from the USSR. NATO would take advantage of the opportunity to extend its front line to the Elbe River, a defensible terrain feature that also would allow NATO to protect a larger portion of German territory.

End Game

For the end game, some form of a flexible, "echeloned" defense doctrine would make most sense. Compared with a linear defense, this doctrine would place fewer forces on NATO's front line and hold a larger portion (e.g., 50 percent) in reserve. Figure S.4 illustrates how an echeloned defense doctrine might be implemented. It envisions a strong frontal array along the Oder-Neisse rivers, and a defense effort that could be conducted

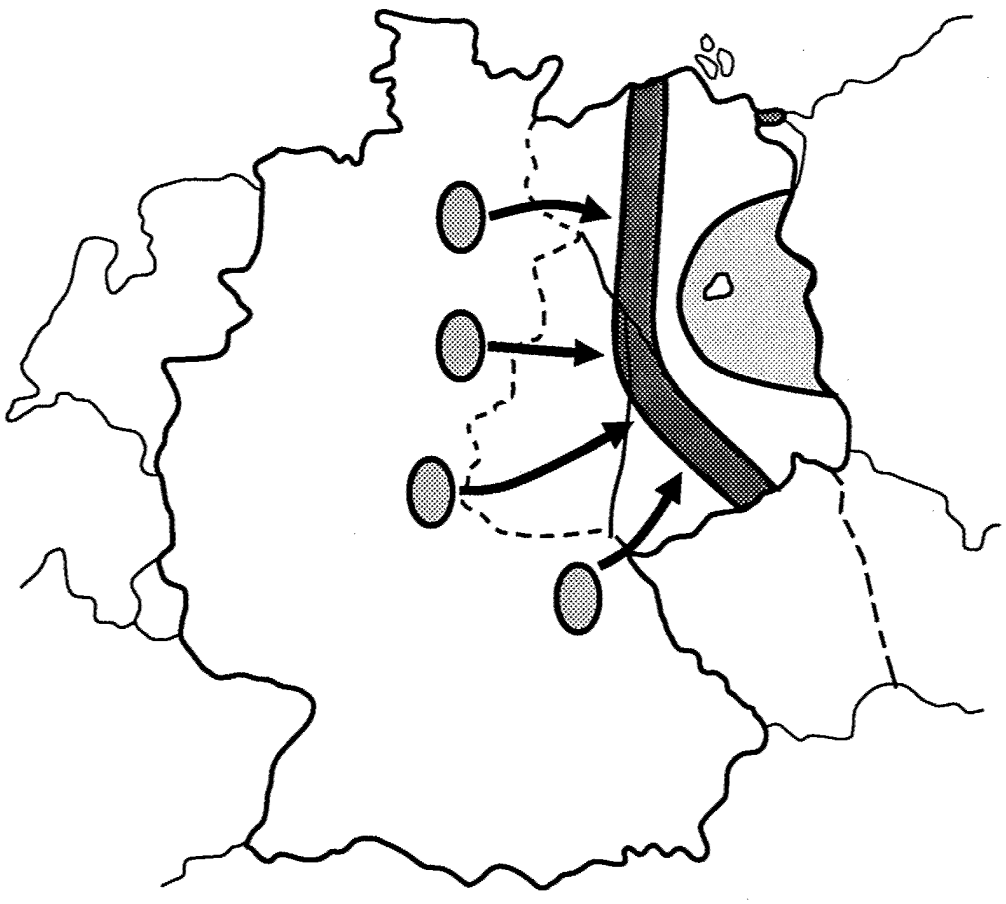

Fig. S.3—Mid-Term: A Transitional Defense Along the Elbe River

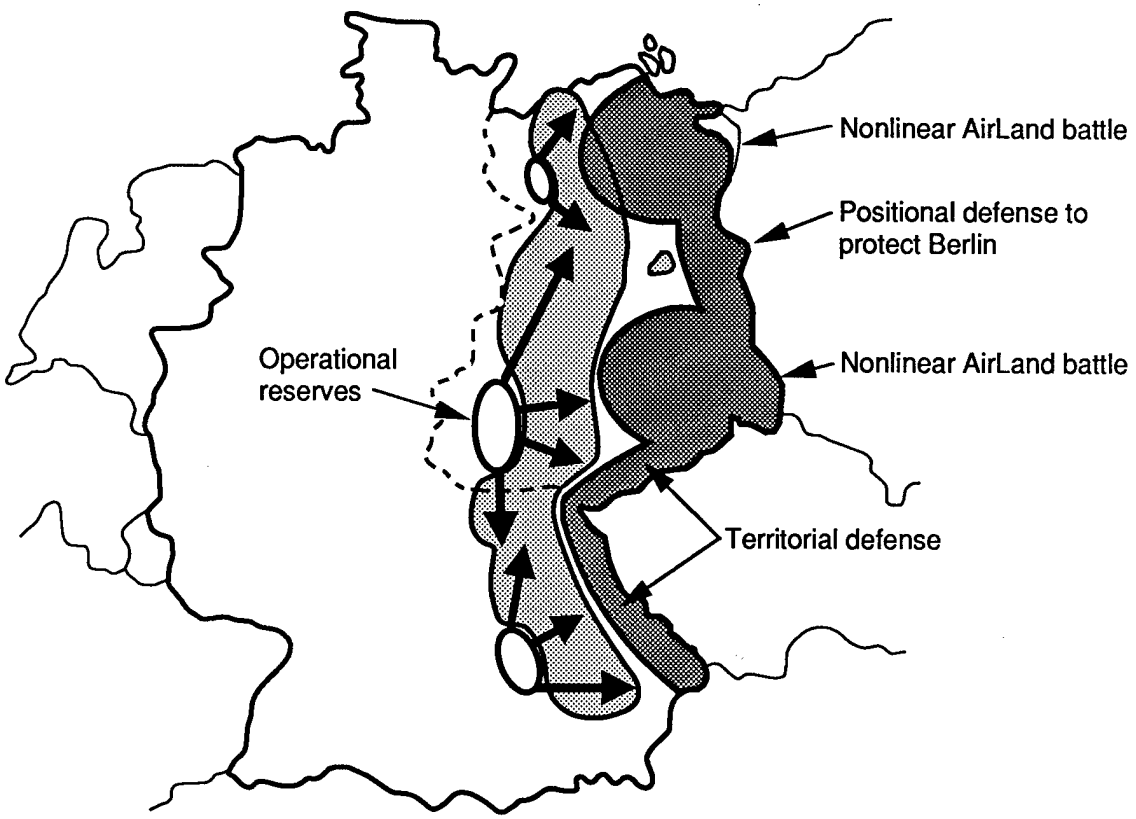

Fig. S.4—Dynamics of an Echeloned Defense

either in a firm way or in more flexible terms. Germany's less-threatened borders along Czechoslovakia would be defended more lightly. Backing up NATO's frontally committed forces would be three large clusters of mobile reserves that could maneuver to meet a main attack. By adopting an echeloned defense, NATO would replace a brittle linear array with a doctrine of frontline forces and large mobile reserves that would be mutually reinforcing.

An echeloned defense of this sort is militarily feasible. Whereas a linear defense of eastern Germany might crumple under the weight of an enemy assault, an echeloned defense—even with NATO forces 10–25 percent smaller than now—should be able to defend successfully. Echeloned defense, however, is not an easy doctrine to master: It would require NATO's forces to conduct coordinated operations in sophisticated ways. Also, NATO would have to make a sufficiently prompt decision to mobilize in a crisis to ensure that its forces had enough time to arrive and to take up their positions. A too-late decision could result in an ineffective defense effort even with a sound doctrine and an otherwise-adequate posture.

A NATO decision to adopt an echeloned defense doctrine would carry with it a need to fashion a replacement for the old layer cake array. Figure S.5 displays one possible rearrangement that would propel NATO in the direction of a "multinational" posture in which the divisions of about 50 percent of NATO's national forces would be brought together into

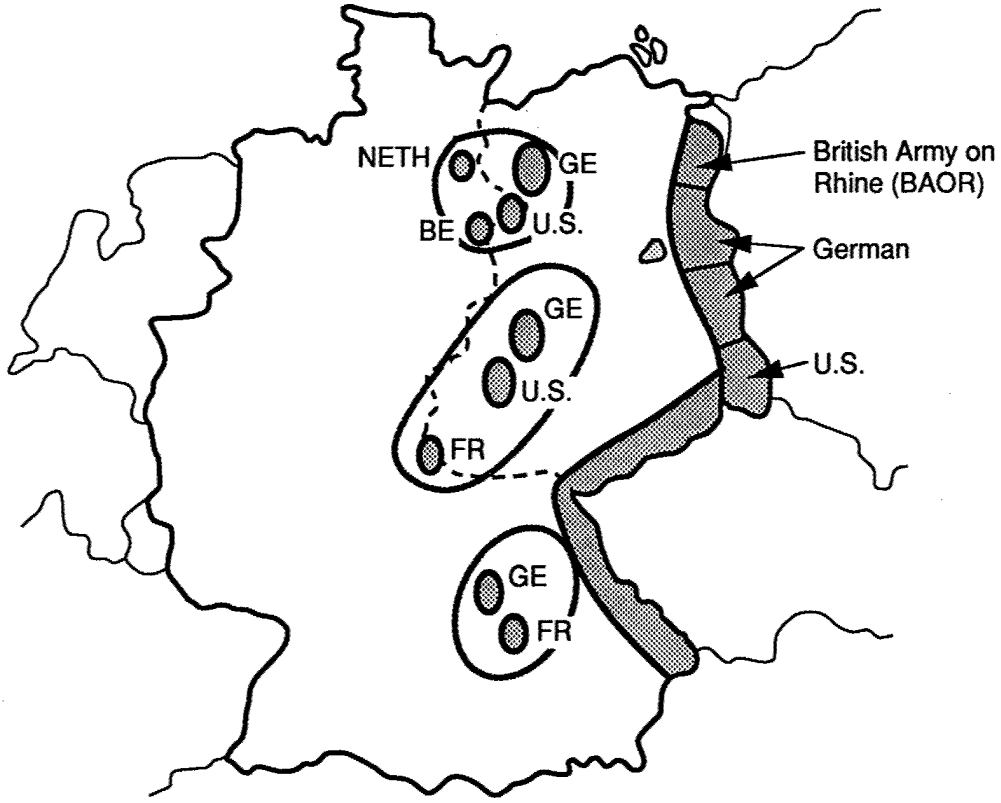

Fig. S.5—A Multinational Approach to an Echeloned Defense Doctrine

combined corps-sized formations. This approach would unambiguously commit NATO to the frontline defense of Germany's new borders. It also would allow NATO to pursue multinational integration. The disadvantage is that it might impede NATO's ability to mobilize quickly in a crisis; also, multinational formations might fight less effectively than national forces.

An alternative to a multinational defense is that of a "specialized" defense. Figure S.6 illustrates how this defense would be conducted. NATO would still be able to form one or two corps of multinational units, but the backbone of its defense doctrine would be divided along national lines. Essentially, the German Army would be tasked with the mission of performing frontal defense roles, while allied forces would provide NATO's mobile reserves. To avoid any singularization of Germany, some Allied brigades would be assigned to frontal roles, and a few German units would deploy as reserves. The U.S. Army, aided primarily by British and French forces, would form the bulk of the reserves.

This approach fosters a natural political division of labor by committing each nation to the task that it seems best suited to perform. Also, it would facilitate NATO's ability to mobilize promptly and fight effectively. The disadvantage is that it might leave Germany concerned about NATO's constancy in a crisis. In any event, it might contribute to a loosening of NATO's coalition bonds, a danger of any specialization scheme.

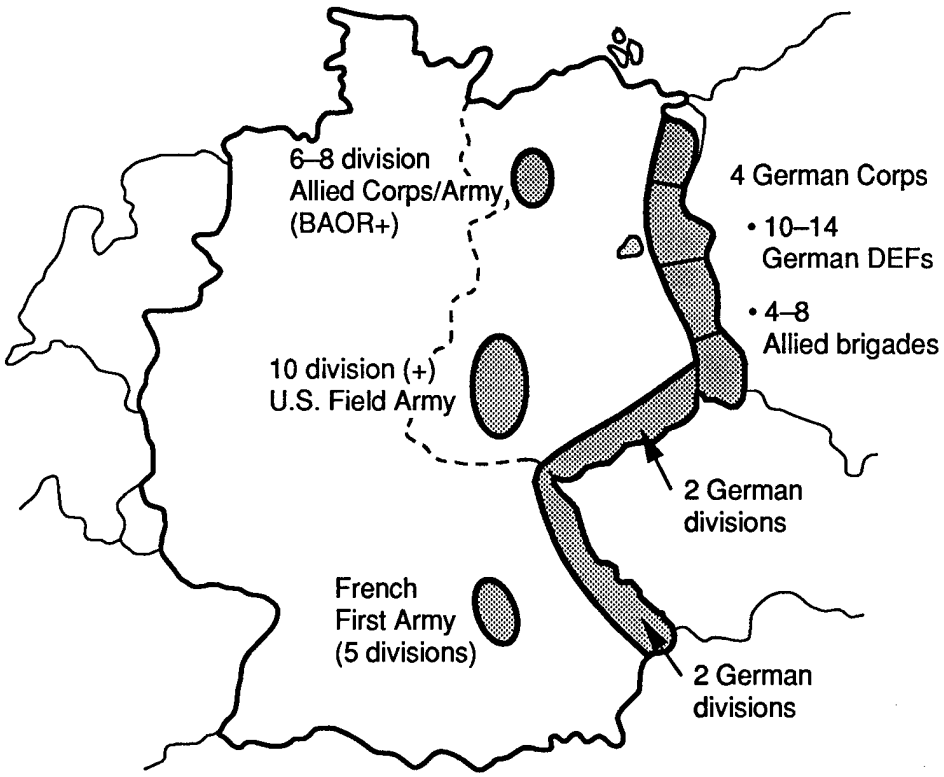

Fig. S.6—A Specialization Approach to Echeloned Defense

NATO'S FUTURE STRATEGY AND FORCE GOALS

These alternatives are illustrations of how NATO might begin thinking about its future theater employment doctrine. They help display the broad range of choice open to NATO, and they permit an initial evaluation of the political-military tradeoffs associated with moving in one direction or another. Above all, they call attention to the need for thorough planning before policy decisions are made. They are not by any means definitive. Detailed analysis will be required of these and other alternatives before NATO selects any future employment doctrine.

These alternatives do suggest an optimistic future for NATO if it crafts an employment doctrine and force posture that respond to the evolving conditions ahead. Optimism is not grounds for euphoria that could lead NATO to disarm too far or otherwise unsafely to relax its military vigilance. NATO's security will still depend on the Alliance's political capacity to mobilize in a crisis. Additionally, even a sound employment doctrine will require adequate forces. A NATO-wide coalition effort will still be necessary. Given its present commitments to force ceilings, Germany cannot defend itself alone against a large enemy force. Even if Germany receives help from its West European neighbors, a sizable U.S. military contribution will be required. The impending withdrawal of about one-half of U.S. forces from Europe therefore should not be translated into an inability to return quickly in a crisis. Provided it avoids these pitfalls by planning intelligently, however, NATO can aspire to meet the

demands of its new conventional military strategy for the coming era. In this important sense, the London Summit was on the right track.

CONTENTS

PREFACE iii

SUMMARY v

FIGURES xvii

TABLES xix

DEFINITIONS xxi

ABBREVIATIONS xxiii

Section

1. INTRODUCTION 1
   - Scope 1
   - Analytical Methodology 3

2. THE DEFENSE ENVIRONMENT IN CENTRAL EUROPE: PAST, PRESENT, AND FUTURE 9
   - Foundations of NATO's Present Conventional Defense Strategy 9
   - Future Security Conditions in Europe 29

3. TOWARD A NEW NATO THEATER EMPLOYMENT DOCTRINE 45
   - Factors Affecting NATO's Decision Calculus 46
   - Alternative Employment Concepts for Defending Along the Old Inter-German Border 52
   - Defending Eastern Germany: The Transition Period 81
   - Defending Eastern Germany: The End Game 89
   - Future U.S. and Allied Military Contributions 112

4. LINEAR DEFENSE IN HISTORICAL PERSPECTIVE 123
   - Legacy of the 19th Century 124
   - Defensive Successes in World War I 127
   - Attacker Successes in World War II 132
   - A Mixed Outcome: The Battle of the Bulge 143
   - Summary Observations 151

5. LINEAR DEFENSE AT LOW FORCE LEVELS: A FORMAL INQUIRY 156
   - A General Model of Linear Defense Against Breakthrough Attacks 158
   - Linear Defense as a Function of Alternative Defender Force Levels 177
   - Sensitivity Analysis 205
   - Summary 211

6. CONCLUSIONS 214

BIBLIOGRAPHY 217

FIGURES

S.1. NATO's Layer Cake Array No Longer Will Be Appropriate viii
S.2. Near Term: A More Flexible Defense of the Old Inter-German Border ix
S.3. Mid-term: A Transitional Defense Along the Elbe River x
S.4. Dynamics of an Echeloned Defense xi
S.5. A Multinational Approach to an Echeloned Defense Doctrine xii
S.6. A Specialization Approach to Echeloned Defense xiii
1. NATO's Layer Cake and Linear Defense 13
2. Historical Trends in WEI/WUV Scores 18
3. Avenues of Advance Along the Inter-German Border 21
4. Deployment Patterns of Soviet and NATO Forces 22
5. Illustrative Campaign Results 26
6. Illustrative IDAHEX Results 27
7. A Future Depolarized European Security System 29
8. Future Soviet Buildup Rates 39
9. Time-Phased Force Ratios in Central Europe 40
10. NATO's Layer Cake Along the Oder-Neisse Line 49
11. Effect of Altering NATO's Defense Line 54
12. Linear and Nonlinear Battlefields Compared 66
13. An Improved NORTHAG Defense Plan 80
14. Elbe Defense Line 82
15. Peacetime Basing Locations of NATO's Ground Forces 86
16. A Southern-Oriented NATO Advance into Eastern Germany 87
17. A Northern-Anchored Advance into Eastern Germany 88
18. The New Dynamics of Military Stability 90
19. Alternative Soviet Buildup Rates in Central Europe 92
20. Military Scenario for a Soviet Attack on Germany 94
21. Adequacy of NATO's Ground Buildup Rate: Early Decision to Mobilize 95
22. Adequacy of NATO's Buildup Rate: Late Decision to Mobilize 96
23. NATO's Tactical Air Forces: Capabilities vs. Requirements 97
24. Soviet Operational Concepts 101
25. Defense in Depth Has Problems 103
26. Mobile Warfare Also Has Drawbacks 104
27. Operational Concept for a Three-Tiered Echeloned Defense 106
28. Future Operational Concepts for Forward Defense 108
29. Operational Concept for a Two-Tiered Echeloned Defense 109
30. A Multinational Approach for Building an Echeloned Defense 113
31. Specialization as an Approach to Building an Echeloned Defense 117
32. Trench Warfare in World War I, 1914–1918 129
33. German and Allied Force Deployments, May 12, 1940 135
34. Barbarossa, June–December 1941 140
35. Allied Defensive Employment Strategy 150
36. Effect of German Unification on NATO's Military Geography 157
37. Relationship Between Defender Force Levels and Defensible Thresholds 159
38. Combat Dynamics over Time with Lanchester Square 167
39. Maneuver Strategy: Single vs. Double Envelopment 173
40. A Typical Breakthrough Battle 179
41. Traditional Soviet/Warsaw Pact Five-Front Attack Formation 186

42. Soviet Front Attack Formation 186
43. Geometry of the New Battlefield 190
44. Illustrative Deployments on D-Day 193

TABLES

1. NATO's Ground Posture, M+30, 1989 8
2. NATO Ground Force Levels in AFCENT 24
3. Estimated Soviet/Warsaw Pact Force Levels in Central Europe 35
4. Soviet Forces vs. Central Region 36
5. Combined Effect of Reductions on Soviet/Warsaw Pact Posture 37
6. An Estimate of NATO's Future Buildup Rates in Central Europe 48
7. NATO/Soviet Buildup Rates 52
8. Effect of Preferential NATO Force Reductions 53
9. Effect of Transferring Forces to NORTHAG 54
10. Effect on NORTHAG's Buildup Rate in Main Sector 55
11. NATO/Soviet Force Ratios in NORTHAG 63
12. Employment Alternatives for a NATO Forward Corps 65
13. NATO Force Needs for a Manageable Force Ratio 111
14. Allied and German Deployments on Eve of the Battle of the Bulge 145
15. U.S. Army Frontages in the Ardennes, December 16, 1944 146
16. Allied Reserve Buildup Rates 154
17. Allied Force Deployments vs. a Real Linear Array 155
18. Employment Pattern of a Linear Defense with 30 Divisions 161
19. Broad Front Attack Strategy vs. Linear Defense 164
20. Effect of Defensive Preparations 166
21. Effect of Defender Preparations on Combat Outcomes on D+5 166
22. Lanchester Square 167
23. Theaterwide Effect of Attacker Concentration Tactics 170
24. Force Deployments vs. the Defender's Army Group A 171
25. Force Commitment Profile for a Single Envelopment Breakthrough Attack 174
26. Range of Outcomes of a Maneuver Battle 176
27. NATO Ground Force Needs According to Traditional Planning Factors 183
28. Soviet Theater Employment Strategy at Low Force Levels: Base Case 187
29. Soviet/NATO Force Ratios 188
30. Illustrative NATO Force Allocations 190
31. Soviet/NATO Force Ratios for Breakthrough Battle 193
32. Army Group A's Force Buildup Rate in the Main Sector 196
33. Effectiveness of Army Group A's Buildup Rate 197
34. NATO's Capacity to Cover an Expanding Front 198
35. Key Features of NATO's Defense Prospects at Alternative Force Levels 200
36. Days of Combat Before Soviet Breakthrough Can Be Achieved 202
37. Postbreakthrough Force Ratios in Main Sector 203
38. Effect of Force Reductions on NATO's Buildup Rate 206
39. NATO's Defense Prospects as a Function of Employment Doctrine 212

DEFINITIONS

Forward Defense. Aimed at defending NATO's borders by engaging an enemy attack as close to those borders as practicable. This term implies no single theater employment doctrine. In theory, it can be implemented with a variety of doctrines.

Linear Defense. Implements forward defense by constructing a thick frontal wall along NATO's borders. So many ground forces would be committed to the forward corps sectors that few are withheld as theater operational reserves (e.g., 10–25 percent or less). For planning purposes, force needs are calculated on the assumption that this defense effort will be conducted in fairly static, positional terms. Tactical maneuvers are executed, but theater-wide counter-concentration maneuvers are not planned for forces originally assigned frontal roles. Firepower, rather than operational-level maneuver, is the primary mechanism by which this employment doctrine is to be executed.

Echeloned Defense. A form of forward defense that constructs a strong frontal wall but also withholds a sizable portion of group forces as a theater reserve. This employment doctrine relies on a combination of firepower and operational-level maneuvers.

Mobile Defense. Withholds most ground forces as an operational reserve. For example, 25 percent of forces might be committed along the front line and the remaining 75 percent used as theater reserves. This concept relies heavily on operational maneuvers (e.g., flanking counterattacks or "hammer and anvil" tactics) as its principal defense mechanism. Although this concept normally is associated with tactics that allow the enemy to penetrate into friendly territory to create opportunities for counterattacks, it can be employed as a form of forward defense if the counterattacks are conducted close to friendly borders.

Nonlinear Defense. A new employment doctrine, popular in U.S. Army circles, that implements a forward defense through a combination of a thin front wall, large mobile reserves, and deep air and missile strikes against the enemy's rear areas and reserve echelons. It aims to defeat the enemy through a combination of firepower and maneuver.

Defensive Defense. An employment doctrine developed in Europe that aims to defend NATO's borders while reducing NATO's offensive potential by removing tanks, self-propelled artillery, and armored fighting vehicles. Essentially, it would construct a thick front wall of infantry forces, with substantial antiarmor weapons based in fixed, static positions.

Layer Cake Array. NATO's traditional plan for executing a linear defense along the inter-German border, composed of eight national corps sectors lined up abreast and organized into two Army Groups.

ABBREVIATIONS

AD	Alliance Defense
AFCENT	Allied Forces, Central
ATACMS	Army Tactical Missile System
ATAF	Allied Tactical Air Force
ATTU	Atlantic to the Urals
BAOR	British Army on the Rhine
C3I	Command, control, communications, and intelligence
CDU	Christian Democratic Union
CENTAG	Central Army Group
CFE I and II	Conventional Forces in Europe (Negotiations)
CFG	Central Group of Forces (Soviet)
CONUS	Continental United States
CSCE	Conference on Security and Cooperation in Europe
DE	Division equivalent
DEF	Division equivalent, firepower
DEM	Division equivalent, manpower
Div	Division
ED	Equivalent division
FEBA	Forward Edge of the Battle Area
FLOT	Forward line of own troops
FOFA	Follow-on force attack
FRG	Federal Republic of Germany
GDP	General defense position
GDR	German Democratic Republic
GSFG	Group of Soviet Forces, Germany
HNS	Host nation support
JSTARS	Joint Surveillance Target Attack Radar System
MD	Military district (Soviet)
NGF	Northern Group of Forces (Soviet)
NORTHAG	Northern Army Group
NSWP	Non-Soviet Warsaw Pact
POMCUS	Prepositioned Overseas Materiel Configured in Unit Sets
RSAS	RAND Strategy Assessment System
SACEUR	Supreme Allied Commander, Europe
START	Strategic Arms Reduction Treaty
WEI/WUV	Weapons effectiveness indices/weighted unit values
WEU	West European Union

1. INTRODUCTION

Dwight Eisenhower once said that plans are nothing, but planning is everything. Experience had taught him that specific plans often had to be tossed into the wastebasket at the moment of truth. But he also believed that the act of making plans beforehand was critical to any military enterprise, in war or peace. By matching ends with means, plans spelled the difference between failure and success.

This study takes Eisenhower's faith in planning to heart. In analyzing the important issue of NATO's future conventional military strategy in Central Europe, it conducts an exercise in planning of the most uncertain sort. Europe is changing so rapidly that any attempt to foresee the future, let alone prescribe how the West should act in it, is a hazardous endeavor. Yet the West cannot afford to let events take their own course, especially in security affairs. To the extent possible, it will need to design its own vision of the future. It will need to guide Europe's evolution in safe directions, and to do this it will need to plan. This study offers concrete ideas that might prove helpful to this effort, but its purpose is more fundamental. It calls attention to the West's need to think seriously about its plans for European defense policy in the coming post–Cold War era.

SCOPE

This study's genesis is the landmark July 1990 NATO Summit in London, where NATO's political leaders pronounced that the Cold War had ended and NATO would become a less military, more political alliance. They also said that NATO will remain concerned about defense problems and militarily vigilant. To this end, they announced that NATO will prepare a new military strategy, one that responds to the major changes underway in Europe. In particular, they said, nuclear forces are to become weapons of "last resort." In adopting this stance, they called for a major departure from NATO's strategy of flexible response, which has long relied on the threat of nuclear escalation as the cornerstone of deterrence.[1]

By deemphasizing escalation, the Alliance's leaders implicitly called for a new NATO military strategy that relies heavily on conventional defense. To this end, they issued broad guidelines on how NATO's conventional plans are to be designed in the years ahead. NATO's new strategy, their guidelines said, should move away from its long-standing emphasis on forward defense and toward "forward presence." In the future, NATO will field smaller, restructured active forces that are maintained at lower readiness, exercise less vigorously, and are composed increasingly of multinational corps made up of national units. These forces, however, are to be mobile, versatile, and capable of reacting to any crisis. Moreover, NATO's active forces are to be backed up by the ability to mobilize larger forces if and when they are needed. In effect, these guidelines call for a conventional posture that is flexible, reconstitutable, and stalwart in its capacity to handle nonnuclear aggression entirely on its own.

---

[1]See London Summit's "Final Declaration," reprinted in The *New York Times*, "Text of the Declaration after the NATO Talks," July 7, 1990, p. 5.

The London Summit thus launched NATO on the path of designing a new and quite different military strategy for the post–Cold War era. For all its innovative qualities, however, the London Summit left much unsaid about NATO's future conventional defense plans and forces. Until its guidelines are elaborated upon, NATO will be hard-pressed to determine what portion of its present force posture it should retain. And until it decides how many forces to keep, it will not be able to define how far it can afford to reduce. Important budgetary decisions, as well as NATO's future military strategy, ride on how the London Summit is interpreted.

This study endeavors to add meat to the London Summit's bones by viewing NATO's future conventional defense plans in Central Europe through the conceptual lens of its theater employment doctrine. "Theater employment doctrine" means the manner in which military force is applied on the battlefield. Shaped primarily by NATO's political-military goals, it determines where NATO's forces would defend and how they would wage coalition war. It decides how forces from several different nations would draw together to fight for a common cause. It also determines how ground forces would be divided between frontline and reserve roles, how frontal and rear-area defenses would work together, and how tactical air operations would be conducted. More encompassing than tactics and operational art, which establish specific guidance for combat, its scope is quite broad.[2]

Theater employment doctrine, in essence, is the critical connecting link between NATO's military strategy and operational concepts. It affects how military responsibilities are allocated among NATO's members and how political influence is distributed within the Alliance. In important ways that often escape public notice, it drives decisions on NATO's forces, programs, budgets, and policies for burdensharing. It is the ingredient that brings NATO's military plans to life and determines whether, in fact, NATO could defend itself in an actual war. Consequently, it will play a large role in affecting how the Alliance's conventional defense strategy is to take shape in the post–Cold War era.

This study contends that the changes at work in Europe will compel NATO to alter its present employment doctrine in major ways. NATO will have to mobilize differently and fight differently. The need to fashion a new doctrine will require NATO to uproot its traditional defense plans in a manner that will affect it not only militarily, but also politically. NATO will need to adopt a new approach to coalition warfare, one very different from that of the past.

In an era when defense issues will command less public attention, NATO will face a serious challenge. In the years ahead, deterrence and defense no longer will be NATO's main preoccupation. But NATO will need to remain militarily strong if it is to maintain a stable balance of power in Europe and underwrite a sound security architecture there. Beyond that, NATO needs a viable military strategy to preserve its own cohesion. In its absence, NATO could come politically unglued; in particular, both the United States and Germany might drift away. Neither development would be healthy for the West's security or for Europe's stability.

The withdrawal of Soviet forces from Eastern Europe will dramatically reduce the military threat that NATO faced during the Cold War. This change will enable NATO to relax its military guard in Central Europe and to view its defense posture as an insurance policy rather than as the foundation of its security. The Soviet Union, nevertheless, will probably remain a military superpower, and NATO will need to guard against any future crisis that could draw it into conflict with that nation. This will remain the case as long as the USSR

[2]For example, a linear defense at the inter-German border is one theater employment doctrine; a mobile defense, either there or elsewhere, is another doctrine.

remains a cohesive nation-state with a nondemocratic government guided by foreign policy goals that potentially run counter to the West's vital interests.

The task of maintaining a conventional counterweight to Soviet power will be complicated by two parallel developments. First, reductions in its own forces will prevent NATO from mounting the linear defense and layer cake array that have been the foundation of its theater employment doctrine for many years. Second, Germany's unification will compel NATO to extend its security umbrella over the former GDR. NATO's military posture will remain based in western Germany in peacetime, but it no longer will defend at the old inter-German border in wartime. It will need to become capable of protecting eastern Germany in any crisis that might bring Soviet forces across Poland and to do so with forces that are too small to conduct a linear defense along old lines.

The following pages discuss these issues in detail. Section 2 analyzes how the political-military changes underway in Europe are compelling a reappraisal of both the conventional balance there and NATO's linear employment doctrine. Section 3, which presents the heart of this study's policy assessment, identifies and examines a spectrum of new theater employment doctrines that could help NATO solve its problems. These alternatives are analyzed in the emerging political situation in Central Europe, but they have broader applicability to other regions and should also be viewed in that context. As applied to Central Europe, each of these alternatives has its own assets and liabilities. This study endeavors to analyze their tradeoffs and implications in a fashion that permits a balanced and well-informed appraisal.

While Secs. 2 and 3 are addressed to policy-oriented readers who are broadly interested in NATO's employment doctrine, Secs. 4 and 5, which provide additional analyses, speak to the concerns of a more technical audience. Section 4 conducts a historical analysis of why, and in what circumstances, linear defense has become vulnerable to attack and defeat; it surveys several important European battles of the 19th and 20th centuries. Section 5 conducts a formal and quantitative appraisal on linear defense's problems and prospects at low force levels; it endeavors to synthesize attrition dynamics, commonly modeled through Lanchester equations, with the complex movement dynamics that take place when a breakthrough attack is mounted against a linear defense. The conclusions developed in these two sections support the policy and strategy analyses presented in Sec. 2, but they also have a broader relevance to a much larger class of defense problems.

ANALYTICAL METHODOLOGY

This study takes aim at a specific policy issue that will be on the European security agenda in the years ahead, but it also endeavors to make a more enduring contribution to the scholarly literature on the theory of military strategy and coalition planning, combining theory and practice. The ideal goal of any scholarly inquiry, theoretical or practical, is to create scientifically valid results, but this state of affairs is seldom approximated, much less attained. A more realistic aim is to forge a useful analysis that, while not removing the need for judgment in decisionmaking, does help educate and inform.

This study presents a multidisciplinary political-military analysis. It examines both the large political changes at work in Europe and the more detailed military transformations that are accompanying these changes. It adopts this approach because the issues raised here are too complicated to be addressed by any single methodology standing alone. A multidisciplinary inquiry is less neat and tidy and more intellectually demanding than a study relying on only one technique. However, it expands the conceptual power of analysis, thereby bringing this study's methodological tools into alignment with the issues raised.

Traditional and formal techniques are used to examine NATO's history in particular and modern military history more generally. Neither the strengths nor the weaknesses of alternative employment doctrines can be understood outside of their historical context, nor can these alternatives be understood outside of their operational context. Therefore, this study also conducts a classical military appraisal. In the manner that professional military staffs might study the issues, it takes into consideration the Central Region terrain, modern military forces and weapon systems, and the battlefield dynamics associated with contemporary military campaign plans.

On the formal side, two analytical methodologies—systems analysis and operations research—help focus the inquiry, provide greater precision, and permit modeling and quantification. As E. S. Quade has said, these two methods differ largely in their analytical scope.[3] Systems analysis is a common technique for analyzing problems of optimal choice when multiple objectives are being pursued and these objectives are either in conflict or at least not easily compared. By taking into account each alternative's consequences in several different areas, systems analysis helps make judgments about the tradeoffs, including costs and benefits. Operations research, in contrast, involves the use of formal models and related quantitative techniques for studying problems of efficient choice. In defense analysis, it is most commonly used when the task at hand is to determine how military forces, operating in a complex operational environment, can best be employed to achieve a given, clearly defined end.

Both techniques must be used here. Systems analysis is necessary because the task of analyzing NATO's theater employment doctrine involves the evaluation of multiple objectives and tradeoffs. More is involved than simply determining which doctrine appears more effective in purely military terms. Other, mainly political objectives must be considered as well. Military goals and complex operational factors also figure heavily. The techniques of operations research must be employed to help understand how modern battlefield dynamics determine the viability and feasibility of each doctrinal option.

This study endeavors to bring the disciplines of systems analysis and operations research together on behalf of a common enterprise. The use of both disciplines is enriched by historical and classical military analyses that together provide both context and a firm anchoring in reality. It is hoped that the outcome will be a set of insights that stand up analytically and make intuitive sense.

Systems Analysis: Criteria of Evaluation

The phenomenon of multiple objectives under conditions of incommensurability implies a complex decision calculus in which costs and benefits are not easily measured and tradeoffs often must be made. Most fundamentally, it means that without some integrating multiattribute utility function, no criterion of evaluation can be used as the only measure of merit. Since these conditions apply here, the systems analysis component of this study employs several criteria that reflect the multiple objectives NATO's defense strategy has pursued in the past and is likely to pursue in the future.

In the examination of linear defense and its alternatives, military stability is a particularly important criterion. This concept assumes that deterrence, one of NATO's highest priority goals, must be met if a particular employment doctrine is to be judged attractive. It goes beyond deterrence to postulate that another important NATO objective—defense—must

[3]See Quade, 1962, pp. 23–24. Also see Keeney and Raiffa, 1976.

also be met. Deterrence is a psychological phenomenon attempting to convince a potential aggressor that the costs and risks of an attack are far in excess of any rational expectation of gain. It could be achieved by forces that bluff an aggressor into believing that the costs of an invasion will be higher than they really are. By contrast, defense—the actual capacity to halt an attack if deterrence fails and aggression occurs—is more demanding. It is anchored on actual military forces and capabilities, and it is measured by physical phenomena. It is achieved only when NATO's forces pass demanding tests of sufficiency that suggest effective performance on the battlefield and thus remove any major incentives for aggression, even to an attacker prone to accepting risks.

The criterion of military stability is derived from the defense objective. It postulates that a primary NATO goal is to create a military situation in Central Europe in which both sides have a confident capability to defend and neither has a capacity to attack successfully. At a minimum, it means that the Soviets should enjoy no confidence in their ability to overpower NATO's defenses in an actual war. Beyond this, it postulates that NATO's forces ideally should provide the Alliance high confidence in its own defensive capability. This means that NATO's conventional defenses should be stalwart and robust. They should not be brittle or prone to fracture and collapse in the event of unexpected and unfavorable turns of events.[4] The criterion of military stability thus carries with it a satisficing standard and a maximizing one. Both need to be kept in mind in evaluating how military stability can best be achieved.

Since NATO's defense planning is driven by other objectives beyond deterrence and defense, this study employs additional criteria of evaluation, including:[5]

- *Forward Defense*, as measured in terms of NATO's ability not only to stop an enemy conventional attack in Central Europe, but also to halt it without ceding large amounts of Allied territory.
- *Control of Escalation*, as measured in terms of NATO's ability to prevent or at least delay nuclear escalation and to conduct nuclear escalation in a controlled and deliberate manner if the nuclear threshold is crossed.
- *Political Stability, Arms Race Stability, and Crisis Management*. These three criteria stem from political objectives in NATO's defense and security planning. The first two criteria refer to the influence of NATO's defense posture on East-West political relations and the military competition in Europe. The third criterion refers to the effect of NATO's military posture on the Alliance's ability to control any European crisis that might occur.
- *Alliance Unity*, as measured in terms of how NATO's defense posture and plans affect the Alliance's internal political cohesion on both sides of the Atlantic.
- *Affordability*, as measured in terms of NATO's ability to implement alternative defense plans with the resources that are likely to be available.
- *Achievability*, as measured in terms of NATO's political and organizational capacity to undertake the kinds of coalition planning and combined operations needed to implement alternative defense plans.

These multiple criteria imply that the task of designing future NATO employment concepts in Central Europe is likely to be complex and demanding. Military stability is clearly a

[4]For a theoretical discussion of military stability as a planning concept, see Rohn, 1990.

[5]For a more detailed discussion of these objectives, see Kugler, 1990b.

dominant consideration, as are the deterrence and defense goals that this concept reflects. But more is involved than military stability alone. Other objectives and criteria must be taken into account as well. Alternatives that are optimal for one set of objectives might appear less attractive when other goals are considered. The task of systems analysis is to identify the tradeoffs involved among alternative employment concepts and to form a balanced appraisal that suggests how each alternative measures up in these several different areas.

Tools of Operations Research

Rather than rely on any single tool, the operations research part of this study employs a family of techniques for gauging how each theater employment concept might perform on the battlefield. The tools used include:

- *Static Planning Factors*. These are "rule of thumb" standards commonly used to help gauge the physical composition of employment doctrines and the conditions under which each doctrine becomes necessary or feasible. They are derived from history, professional military judgment, and classical military methods. Their weakness is a lack of dynamic properties; their strength is their clarity and transparency.
- *Simple Dynamic Models*. Employing transparent algorithms for allocating forces on the battlefield, these "back of the envelope" models are useful for assessing how different force allocation strategies can shape the battlefield in ways that alter the local balance of forces. These models are responsive to some important dynamics and are transparent; their drawback is that they cannot provide the depth of dynamic computer models.
- *Dynamic Computer-Based Simulation Models and War-Gaming Systems*. This study will employ a combination of models. Their strength lies in their capacity to measure how multiple variables interact and change over time, thereby assessing the important dynamic properties of war. Their weakness lies in their internal complexity, their uncertain capacity to capture operational realities, and their lack of transparency.

Sensitivity Analysis

Like defense planning itself, operations research is conducted in the face of uncertainty. To establish boundaries around the analytical process, assumptions have to be made about how the forces on both sides would perform in a war. This study is no exception. Its assumptions (and supporting data) broadly reflect previous studies on the Central Region force balance done by RAND and the Department of Defense. Its analytical thrust is mainstream.

Regardless of the extent to which analysts may have accepted any set of assumptions, an actual war could unfold in quite different ways than that set forecasts. The defense equation, after all, is dominated by variables, not constants, and it is influenced by exogenous factors and even intangibles that lie beyond the boundaries of normal inquiry. A critical issue in any analysis therefore is how much the study's principal conclusions are sensitive to variations in its underlying assumptions, explicit and implicit. Recognizing that, this study employs sensitivity analysis to determine whether its major judgments about NATO's doctrinal alternatives are robust or brittle: whether they stand up strongly to changes in assumptions or break down readily.

In employing sensitivity analysis for this purpose, this study distinguishes between "expected value" and "prudently conservative" assumptions. The difference between these two categories is important and needs to be treated explicitly. "Expected value" assumptions manage uncertainty by postulating that war is likely to unfold along the most probable lines. They thus exclude the improbable, but still possible, ends of the spectrum. Because they presume no improbable conditions for either side, these assumptions place no more stress on NATO's posture than is most likely to occur in wartime. They most commonly are associated with a risk-neutral policy that adopts an indifferent, or "break even," stance to uncertainty and that does not seek a margin of insurance against the unlikely.[6]

By contrast, "prudently conservative" assumptions reflect events that are less likely to occur but still lie well within the range of plausibility. Pearl Harbor, for example, was such an event. Prudently conservative assumptions are not "worst case" in nature, but they intentionally place greater stress on NATO's posture by reflecting conditions that seem unusually favorable to the enemy. They normally are associated with risk-averse policies of the sort favored by many defense planners that seek assurance against improbable but costly events.

In past years, U.S. defense policy often has been formulated under risk-averse principles; indeed, "prudent conservatism" has come to be accepted as a hallmark of sound planning in many quarters. Whether this approach, with its unique set of strengths and drawbacks, will continue to be employed in the future is uncertain. Regardless, it is important to know whether NATO doctrinal alternatives that appear attractive under expected value conditions retain their luster when prudently conservative assumptions are employed, and vice-versa.

Static Force Measurement Techniques

Finally, this study employs a family of static measures for characterizing the strength of NATO and Soviet/Warsaw Pact forces in Central Europe. These measures do not differ radically from each other, but they are not analytically identical, and their differences occasionally do matter. Consequently, each of them should be used in an internally consistent fashion and with regard to the circumstances immediately at hand. They are:

- *Division* (Div). A division is defined as a ground force maneuver formation that is formally labeled (flagged) as such and totals about 12,000–16,000 soldiers. By traditional Western standards, a division has been regarded as ideally defending a frontage of 25–30 kilometers (km) and attacking on a minimum frontage of 10–15 km. Sensitivities will be addressed later.
- *Division Equivalent* (DE). A division-equivalent is either a division-flagged formation or a hypothetical combination of three independent brigades or regiments whose total size and combat power is roughly equal to a standard division. For example, a NATO posture of 40 divisions and 15 independent brigades/regiments would total 45 division-equivalents.
- *Division Equivalent, Manpower* (DEM). This term measures total combat manpower rather than numbers of maneuver formations. A DEM is composed of 16,500 soldiers, the size of a U.S. Army or German division. A NATO posture of 720,000 soldiers would have 45 DEMs. A posture with units of 561,000 men would have 34 DEMs.

[6]For a theoretical discussion of risk and uncertainty in decision models, see Keeney and Raiffa, 1976, pp. 66–353.

- *Division Equivalent, Firepower* (DEF). The DEM measures manpower, not weapons, and thus does not gauge a formation's total firepower. The DEF is a firepower measure that counts the weaponry of ground combat formations, including corps artillery and attack helicopters. Based on quantitative and qualitative factors, it employs the WEI/WUV III system.[7] A single DEF is equivalent to the WEI/WUV score of a standard German heavy division (the average of German armored and mechanized divisions), about 80,000 WEI/WUV points. Thus, a NATO posture of 45 DEs that generated 4.5 million WEI/WUV points would receive a score of 45 DEFs. But this coincidence of numbers is not automatic. The WEI/WUV system places heavy emphasis on armor and artillery and less on infantry-associated weapons. Therefore, a posture composed of infantry units (each amounting to only .5 DEFs because of fewer heavy weapons) would total only 22.5 DEFs. Also, a total score of 45 DEFs does not mean that each division formation fields a 1.0 DEF score. Major variations can, and do, occur among NATO's national armies; U.S. Army units have scores above the 1.0 standard, while Belgian and French division scores are much less.
- *Equivalent Division* (ED). The ED, which is used in several other RAND studies, is a firepower measure that also uses the WEI/WUV III system. Its standard of comparison, however, is based on the score of a typical U.S. Army heavy division with modern equipment (about 100,000 points), which is higher than a standard German division. For this reason, a NATO posture that totals 45 DEFs would amount to only 36 EDs.

This study relies most heavily on the DEF measure; the other techniques are used in a supplementary way. Table 1 helps clear up the ambiguity surrounding these static techniques and provides the reason why the DEF measure is most heavily used in this study. An estimate of NATO's 1989 posture in Central Europe after about one month of mobilization shows that the DEF measure is similar to the DEM and DE and thus is a satisfactory measure of total force size and mass: NATO's strength in weapons and manpower is the equal of about 45 German divisions. The DEF also measures NATO's total firepower in weapons; like the ED, it can be used to make comparisons with Soviet/Warsaw Pact forces. If one measure is to be used, the DEF, for the purposes of this study, is probably the best available.

Table 1

NATO's Ground Posture, M+30, 1989

Divisions	41
DEs	46
DEMs	44
DEFs	45
EDs	36

[7]See U.S. Army Concepts Analysis Agency, 1978.

2. THE DEFENSE ENVIRONMENT IN CENTRAL EUROPE: PAST, PRESENT, AND FUTURE

Although it is impossible to foretell exactly how Europe will evolve in the coming years, the future military situation in Central Europe will be quite different from that in the past. This section sets the stage for analyzing NATO's future theater employment doctrine by examining how the changes at work there will affect NATO. It reviews NATO's Cold War military strategy for managing the NATO-Warsaw Pact confrontation at the old inter-German border, discusses how impending force reductions on both sides and the dissolution of the Warsaw Pact will transform the military balance in Central Europe in less threatening but still challenging directions, and examines how Germany's unification will extend NATO's military commitments eastward.

FOUNDATIONS OF NATO'S PRESENT CONVENTIONAL DEFENSE STRATEGY

Since 1967, NATO's strategy has been one of "flexible response" (often called MC 14/3, after the title of the NATO document explaining its contents). The London Summit set in motion a review process that will culminate in adoption of a new strategy, presumably to be called MC 14/4. This new strategy will downplay the role of nuclear weapons and elevate the importance of conventional defense. But the degree to which NATO's strategy will change might be less dramatic than surface appearances suggest.

MC 14/3 itself went a considerable distance in this direction. It stopped well short of labeling nuclear forces as weapons of "last resort," but it recognized the dangers of a strategy that relied on escalation for dealing with limited war. Also, it called for a strong conventional deterrent. The London Summit suggests that MC 14/4 will replace MC 14/3's shrill call to arms with a post–Cold War stance envisioning a "regenerable" defense posture, rather than one that could fight at a moment's notice. But in shifting emphasis toward a defense policy of backup insurance, the new strategy will not discard MC 14/3's emphasis on flexibility. The changes that it institutes in this area will be evolutionary, rather than revolutionary.

Major changes will come in the superstructure of policies, plans, and programs that give conventional defense its meaning. If MC 14/3's experience is an indicator, these changes will take some time to define and work out. In any event, theater employment doctrine is one critical part of the military superstructure that will have to be created. Arguably it is the most important part because so many decisions—political and military—cannot be made in its absence.

Before MC 14/3 was adopted, NATO embraced a quite different strategy. From 1950 to 1957, NATO's defense plans were forged under the auspices of MC 14/1, a strategy that called for a large conventional defense posture in Central Europe—54 divisions were envisioned by the 1952 Lisbon goals—backed up by the U.S. arsenal of strategic nuclear bombers. In 1957, this strategy was replaced by MC 14/2. Like its predecessor, MC 14/2 emphasized deterrence as its primary objective. But unlike MC 14/1, the new strategy sharply downplayed conventional preparedness by emphasizing nuclear escalation as its principal deterrent mechanism. Philosophically, it aimed to offset the Warsaw Pact's conventional preponderance in Central Europe by relying heavily on U.S. nuclear superiority.

NATO's conventional defense goals centered on fielding only enough active forces to establish a brief pause before launching a massive U.S. nuclear strike against the Soviet Union.[1]

After a long debate that was instigated by the United States, NATO abandoned MC 14/2 in 1967 and substituted MC 14/3. The strategy of flexible response replaced the plan for almost immediate and massive escalation with a three-tiered strategic concept of direct defense, deliberate escalation, and general nuclear response. NATO retained the flexibility to cross the nuclear threshold quickly if the Warsaw Pact employed nuclear weapons from the outset, but in the event the Soviet Union initially employed only conventional forces, MC 14/3 called on NATO to respond in kind. It envisioned a "direct defense" that would be aimed at denying the USSR its political and military objectives. Should the direct defense fail, MC 14/3 then called for a "deliberate escalation," including the use of nuclear weapons, but the initial employment of these weapons was not to be massive. Instead, escalation was to be conducted in a controlled manner, with the goal of locally halting the enemy's advance and influencing the Soviet government's willingness to continue the attack. Only in the event that this tactic failed, under conditions involving the impending defeat of NATO's forces or a threat to the survival of a NATO nation, was a massive, strategic nuclear strike against the Soviet Union to be launched.[2]

MC 14/3 sought to deter by creating viable defense options that would be militarily effective on the battlefield and politically acceptable to NATO. It endeavored to replace the prospect of premature escalation with a better NATO capacity to control the battlefield and to transfer the onus of escalation onto the shoulders of the aggressor. It retained the option of crossing the nuclear threshold first, if events warranted that step. But its calculus of deterrence, involving the interaction between conventional and nuclear forces, was a good deal more sophisticated than that of MC 14/2.

NATO's adoption of MC 14/3 reflected the fact that the Soviet Union was acquiring a strategic nuclear arsenal of its own. MC 14/3's authors believed that the threat of immediate escalation no longer would be a credible deterrent to the Soviet Union. They worried that the Soviet government, operating under its own nuclear umbrella, might conclude that NATO would shy away from escalating, especially against a limited attack that did not immediately threaten the entire Alliance. In trying to restore credibility to NATO's strategy, MC 14/3 was especially designed to deter limited "salami slice" attacks, as well as unlimited aggression. Additionally, it strove to restore NATO's badly frayed unity by creating a strategy that would safeguard the interests of both the United States and the West Europeans. MC 14/3 thus pursued multiple objectives and faced inward as well as outward.

Integral to MC 14/3 was the goal of building a stronger conventional defense posture in Central Europe as well as on the flanks. MC 14/3 left ambiguous its precise conventional force objectives, but the basic concepts that it established were clear enough. At a minimum, it called for a force posture that could deny the Warsaw Pact a quick and easy advance into NATO's territory. It also aimed to give NATO's governments at least enough time to resolve their own uncertainty before being compelled to escalate. Its ideal, or maximalist, goal was to build a posture strong enough to rebuff an attack without resort to nuclear weapons at all. MC 14/3 thus postulated a floor and a ceiling for NATO's force plans; the floor was to be laid immediately and the ceiling erected over a longer period.

Commitment to forward defense was an important feature of MC 14/3 and especially critical to West Germany, which feared that flexible response might be interpreted as a

[1]For a discussion of MC 14/2, see Kugler, 1990a.

[2]See Kugler, 1990b. See also Stromseth, 1988.

slackening of NATO's resolve to protect frontline states. Driven partly by concern for its many urban areas near exposed borders, the West German government would not have agreed to MC 14/3 had not forward defense been incorporated. The drive to build consensus support for MC 14/3 had a direct bearing on NATO's employment doctrine. In the early 1950s, NATO's plans contemplated a defense as far back as the Rhine River. Later that decade, NATO's defense line was moved forward, but it was still drawn well back of the border, in the vicinity of the Weser-Lech river line. In the 1960s, in response to the wishes of West Germany, the line was moved to the border region itself.

The eastward march of NATO's defense line was driven partly by political imperatives. The idea of forward defense originated in NATO's early days, but not in West Germany, which at that time was neither a sovereign nation nor a member of NATO. The French, Belgian, and Dutch governments were all nervous about a Rhine defense that, if breached, would leave their nations vulnerable. It was they who insisted that NATO attempt to defend "as far forward as possible" well east of the Rhine. West Germany entered the picture in 1955, when it joined NATO. It conditioned its willingness to rearm on NATO's agreement to defend German territory; and although it initially stated its preferences cautiously, it steadily became more vocal. By the early 1960s, forward defense was a concept whose time, politically speaking, had come.

NATO's decision to defend forward was propelled by more than politics; military factors mattered as well. In early 1950, NATO had only nine ground divisions deployed in Central Europe. It was badly outnumbered by the Soviets, unable to cover much terrain, and far too small to defend along the German partition line. The decision to beef up the U.S. military presence, to adopt the Lisbon force goals, and to build a West German Army of 12 divisions began changing things. By the late 1950s, enough forces had been fielded—roughly 20 NATO divisions—to permit adoption of the Weser-Lech concept. Not by accident, NATO's decision to move its defense line further forward coincided with the completion of the German rearmament effort. Thus, NATO's defense line marched eastward in tandem with its growing military strength.

During the 1960s, forward defense was not interpreted rigidly. Allowance was made for flexible employment tactics in which ground would be ceded locally to allow for effective force employment. But neither was forward defense a hollow concept. NATO made a firm commitment to defend forward and not to resort to nuclear escalation until this effort had been irrevocably defeated. From that point onward, NATO's military plans and programs were tailored with that goal in mind. It was the case for the five-year improvement plan adopted in 1967 along with MC 14/3. It was also true for subsequent NATO plans, including AD-70 of the early 1970s, the Long-Term Defense Plan later that decade, and the Conventional Defense Improvement plan of the 1980s.[3]

Linear defense made its appearance in the 1960s, when NATO formed its "layer cake" array, and it has remained NATO's bedrock ever since. The programs of the 1970s and 1980s concentrated on making linear defense and the layer cake viable. By committing most of NATO's forces well forward and along a front line, the linear defense concept marked an important swing away from the idea of mobile warfare, which had figured heavily in NATO's earlier doctrine. The layer cake was NATO's chosen vehicle for implementing its linear defense. It created eight national corps sectors lined up abreast (German and Danish forces

[3]See Kugler, "Surviving Forty: Toward a Robust NATO Conventional Defense Posture," in Golden et al., 1989.

deployed in Jutland formed the equivalent of a ninth corps). Guarding the northern Federal Republic of Germany (FRG) were four corps under the Northern Army Group (NORTHAG). Guarding the southern FRG were four corps under the Central Army Group (CENTAG). Both NORTHAG and CENTAG were placed under command of Allied Forces, Central (AFCENT), which reported to the Supreme Allied Commander, Europe (SACEUR).

Together, these corps covered the entire Central Region border from Schleswig-Holstein in the far north to the Austrian border in the far south. NATO's combat forces themselves were not heavily deployed in the forward areas in peacetime; most remained based at locations somewhat to the rear. NATO's commanders developed plans to move these forces to forward locations once an alert had been called, positioning them to meet an enemy attack on the border. The guiding concept was to quickly establish a strong frontal wall that could not easily be breached anywhere.

Specific force commitments within each corps sector have changed periodically over the past two decades, but the basic structure of the layer cake has remained constant. Figure 1 displays the layer cake as it was envisioned through 1989, along with an illustrative estimate of NATO force commitments in each sector and the forces that would be available after one month of mobilization. What stands out is the extent to which NATO's forces have been planned for commitment into the forward areas. Some 34 of NATO's 45 DEs, 75 percent of the total posture, are deployed into the forward corps sectors. NATO's operational reserves, units withheld under Army Group or AFCENT command, are limited to reinforcing U.S. III Corps and a few British and German units. Added to this list is the French Army, which is not formally committed to NATO but probably would be made available in a crisis. NATO's "center of gravity" thus has been placed well forward.[4]

An important feature of the layer cake is that it has distributed NATO's forces uniformly among the individual corps sectors. It gives each corps 3–5 DEs, counting cavalry troops that conduct reconnaissance. Looking from north to south, the first six corps average about 60 km of frontage (as does the Jutland sector). Divisions actually committed to front-line roles have frontages that average 25 km. This uniform deployment scheme is no accident: It reflects doctrinal norms and an underlying concept aimed at ensuring that each corps could not be dislodged quickly. The final two corps sectors (U.S. VII and German II corps), which cover the heavily wooded terrain along the German-Czechoslovakian border, are much longer and less densely defended. To some extent, their thin configuration reflects the more defensible terrain there. But it also reflects a conscious decision to accept greater risk to concentrate more forces northward where loss of ground could be crippling.

The lack of depth to NATO's posture behind the forward corps sectors would change after a longer period of mobilization as additional U.S. Army units arrived from the United States. By M+75, the deployment of 10–15 U.S. DEs would provide larger reserves to back up the forward corps sectors. Nevertheless, in the early stages, during which many analysts have feared a Soviet attack, the preponderance of NATO's ground power would be deployed near the inter-German border and spread out along a long (750 km) north-south axis. NATO's forward wall would not be, as some critics have alleged, paper thin. It broadly conforms to Western military standards for corps that are intended to defend intensely and have some staying power. But NATO's frontal wall would lack the backup reserves that military commanders traditionally have regarded as necessary for a stalwart defense that can adjust to reversals.

[4]See Phillip Karber and A. Grant Whitley, "The Operational Realm," in Golden et al., 1989. See also Lawrence and Record, 1976; Mako, 1983.

NATO's present forces (D-day) are planned according to these standards

Corps sectors	Division-equivalents	Frontage (km)
Jutland I	3 1/2	60
Neth I	3 1/2	65
FRG I	4 1/2	60
UK I	4	55
BE I	2 1/2	40
FRG III	3 1/2	65
US V	4 1/2	75
US VII	5 1/2	155
FRG II	4 1/2	175
	36	750
Theater reserves	9	
	45	

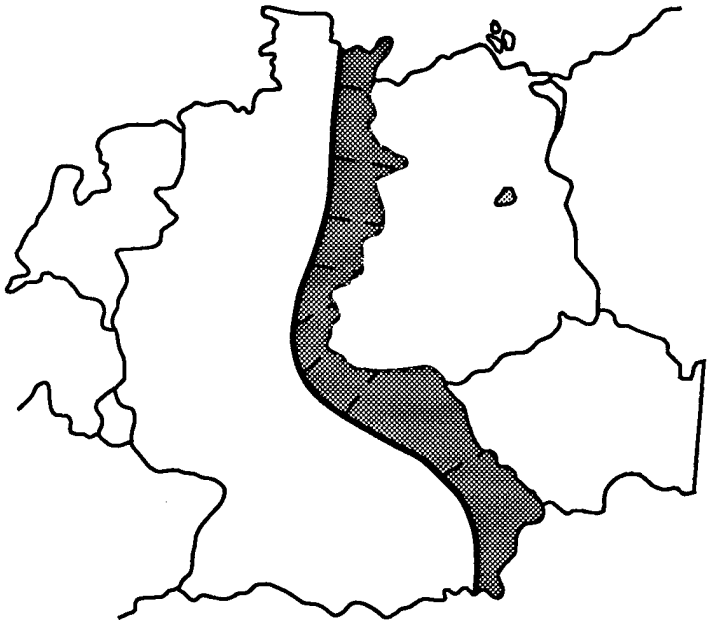

Fig. 1—NATO's Layer Cake and Linear Defense

NATO's reliance on a linear defense and the layer cake do not mean that its military plans have been immobile or rigidly committed to protecting Germany's borders. Over the years, NATO's doctrine for executing the defense has undergone a series of changes driven by shifting theories for interweaving firepower and maneuver. In the 1960s, prevailing tactics were especially flexible in their interpretation of forward defense. Commanders planned to hold several NATO units in the corps rear and to yield ground locally. This tactic, which reflected an emphasis on maneuver rather than firepower, was adopted to channel enemy forces into killing zones and to permit NATO flanking attacks against them. NATO thus had a linear defense and layer cake array but implemented them through a form of mobile warfare.

In the 1970s, the pendulum swung in the other direction as the "active defense" doctrine came into being. This doctrine made NATO's plans more linear, more frontal, and more positional than before. It called for most forces within each corps to be deployed directly on the forward edge of the battle area (FEBA). The goal was to defeat enemy forces with an explosion of firepower to be generated by large NATO forces lined up abreast and firing from behind prepared positions. The active defense was not a totally stationary doctrine, since divisions and brigades within each corps employed mobile tactics as a device to bring their firepower to bear. But it did place NATO's center of gravity far forward, it relegated maneuver to a secondary position, and it planned to yield less ground than its predecessor. Its leitmotif was "target servicing," a phrase that aptly captured its emphasis on mechanically applying firepower to destroy the enemy physically without first unhinging him operationally.[5]

[5]Herbert, 1988.

Another distinguishing feature of the active defense is that it concentrated on building a strong frontal wall, rather than using a frontal wall and deep reserves together. It endeavored to strengthen NATO's forward corps on an individual basis, breaking the theater battle into several discrete parts, all conducted linearly in the frontal zone. It made no special provision for conducting coordinated maneuvers that would enable frontally deployed corps to support each other. Also, its scheme did not include the use of deep reserves in massed formations and in mobile ways. Reserves were to be parcelled out to the forward corps bit by bit. In these ways, the active defense paid little attention to the "operational art," or the capacity to see the theater as a whole and to blend NATO's front and rear corps-level operations together to conduct an integrated theater campaign.

The doctrinal pendulum began swinging back toward maneuver in the early 1980s as the AirLand Battle doctrine entered NATO's thinking. Driving the AirLand Battle was concern that a linear NATO defense would be vulnerable to Soviet echeloned breakthrough tactics in ways that the firepower of frontline forces alone could not offset. It presented a scheme for destroying the enemy's first echelon forces while his second echelon was held at bay. With the prospect that the Army Tactical Missile System (ATACMS) missile and Joint Surveillance Target Attack Radar System (JSTARS) would soon enter NATO's inventory, the AirLand Battle envisioned the use of deep fires into the enemy's rear areas to disrupt the flow of enemy second echelon forces into the frontal battle. NATO's air forces, benefitting from technology advances of their own, were to aid this effort with interdiction strikes. Against the first enemy echelon, NATO's ground forces were to rely less heavily on fixed positions and massed firepower. More corps forces were to be held in reserve, from where they could be used in counterattack roles. The AirLand Battle thus planned to conduct the frontal defense with a richer combination of fire and maneuver than did the active defense doctrine.

What also distinguished AirLand Battle was its greater emphasis on the operational art. Its emergence coincided with U.S. efforts to complete its program to preposition the equipment of an additional three Army divisions in Europe (Prepositioning of Materiel Configured in Unit Sets, or POMCUS). This program was designed to permit the rapid deployment of U.S. III Corps. A decision was made to retain III Corps, composed of three armor-heavy and mobile divisions, as an AFCENT reserve. Emphasis was placed on using this powerful force in the vulnerable NORTHAG area, in support of the frontal wall there. Initially, NATO planned to feed III Corps units into the front line, but eventually decided to use it as a true reserve, with emphasis on flanking attacks against enemy salients. In this way, III Corps helped reintroduce operational art into NATO's doctrine.[6]

Other steps were taken to enhance NATO's deep reserves. In NORTHAG, British commanders developed plans to withhold a few U.K. and German forces in a reserve capacity. The French began developing plans for deploying the rapid-reaction force and their own III Corps in more creative, AFCENT-wide ways. Before this innovation, nearly all French forces had planned to deploy into the CENTAG region, where they could screen the French border. France remained outside of NATO's integrated military command, but its doctrine was paying closer attention to that of NATO, and its forces were trying to become more skilled in operational art.

NATO's defense doctrine of the 1980s thus was not rigid or stationary. It relied on tactical mobility, and it was sensitive to the need to form reserves. NATO's modernization

[6]See Pierre, 1986.

programs, which stressed battlefield mobility and high-speed operations, were specifically aimed at enhancing its capacity to maneuver. Procurement of the M-1 Abrams tank, the Bradley infantry fighting vehicle, and modern attack helicopters enhanced the U.S. Army's capacity to maneuver. The German Army followed a similar path, as did the other allies to a lesser extent. These doctrinal and programmatic changes were an important departure, but they were hardly revolutionary. In many ways, they reflected continuity in NATO's forward defense concept, which, in varying degrees, always has taken mobility and maneuver into account.

Nevertheless, the layer cake remained the backbone of NATO's defense plan, and it continued to operate in predominantly linear and positional ways. Large corps formations still planned to line up abreast of each other, with only about 25 percent of NATO's forces withheld as a reserve. Forces assigned to individual forward corps sectors planned to remain there during the course of battle. Plans were adopted to shift corps reserves to neighboring corps sectors if the tactical situation dictated, but divisions initially committed on the front line did not plan to redeploy in a large-scale way. In other words, NATO's front line planned to remain intact, come what may.

Perhaps most important, NATO's defense plans did not envision heavy reliance on theaterwide redeployments. Apart from U.S. III Corps and some French forces, the AFCENT commander had few units under his direct control. NATO's Central Region posture was divided between NORTHAG and CENTAG. NATO's tactical air forces were capable of promptly swinging from north to south and were planned to be used in this capacity, but NATO's slower-moving ground forces were not similarly oriented: CENTAG forces did not plan to redeploy to NORTHAG, or vice versa. The two Army Groups planned to fight separate battles, drawing on the forces originally committed to them. This approach gave each Army Group sizable forces for conducting the forward defense in its area of responsibility. But in the absence of large outside reinforcement from the United States, each was left susceptible to being overrun, without much help from the other.

NATO's linear defense thus lacked both sizable deep reserves and the capacity to shift large forces from one area to the next. This deficiency created an Achilles heel in Alliance defense strategy: vulnerability that the enemy might concentrate against one Army Group and then prey upon its lack of reserve. Nevertheless, the decision to maintain a linear defense was not made lightly or without serious study of the alternatives. Positive reasons undergirded this approach to deterrence and war-fighting. These reasons need to be understood if post–Cold War alternatives to linear defense are to be evaluated with a clear eye.

Political Reasons for Linear Defense

Compelling political reasons lay behind NATO's linear defense and layer cake array. Put simply, linear defense required a coalition response. The German Army was far too small to provide the 30–35 divisions needed to establish a strong frontal wall of eight large corps sectors. Nor could German and American forces accomplish the task together. Linear defense compelled the British, Dutch, Belgians, and Danes all to make a formal commitment to flexible response and forward defense. Indirectly, it also exerted a constraining force on France's ability to disentangle itself from NATO's military doctrine. Linear defense therefore played a role in giving birth to MC 14/3. In later years, it helped to sustain NATO's commitment to MC 14/3 by giving Allied governments an incentive to maintain adequate forces and programs. It ensured that weak defense programs would directly expose uncooperating

nations, as well as the whole Alliance, to immediate danger. By making forward defense a multilateral responsibility, linear defense and the layer cake helped to forge and cement NATO's support of its military strategy.

A linear defense array also sent a clear and unambiguous signal that the Alliance was firmly committed to defending West Germany's borders. The plan to establish a strong frontal wall had an important salutary effect in West Germany, especially during the 1960s when the FRG was coming to question American constancy and Alliance commitments. Beyond this, the layer cake sent an important deterrent signal to the Soviet Union and the Warsaw Pact. It told the Soviet government, plainly and unmistakably, that NATO was serious about protecting West Germany and that "forward defense" meant what it said. At the time the layer cake was formed, Cold War tensions still ran high in Central Europe, and the possibility of another Berlin crisis loomed ahead. By enabling NATO to quickly construct a military wall of its own on the FRG's borders, the layer cake was intended to help stabilize any future crises that might explode across the East-West partition line.[7]

Linear defense meshed well with NATO's nuclear doctrine and its political imperatives. Although vulnerable to sustained pounding, linear defense helped ensure that NATO's forces would not have readily exploitable vulnerabilities and that attacking Soviet forces would not be able to quickly advance deep into West German territory. It provided NATO the initial period of conventional defense that the United States demanded as a condition for providing extended nuclear deterrence. And it bought NATO the time—at least several days—that all NATO governments agreed was needed for the Alliance to make the momentous decision to cross the nuclear threshold. By design, it helped ensure that Soviet forces could advance into West Germany only if the Soviet government had made a firm decision to go to war and had committed the large Warsaw Pact forces that signalled unlimited intent. Also important, linear defense provided an unambiguous criterion for actually deciding to escalate after NATO had exhausted its conventional options. A major breach of NATO's forward wall potentially enabled SACEUR to argue convincingly that NATO's conventional forces were on the brink of defeat and that the situation could not be salvaged. Moreover, a successful penetration of the layer cake's rear boundaries would have left Soviet forces descending on West Germany's industrial areas (Frankfurt and the Ruhr) near the Rhine River. Under these adverse political and military conditions, escalation would have become the kind of alternative that all NATO governments plausibly could perceive as necessary.

At the same time, linear defense helped ensure that if nuclear escalation became necessary, initially it could be conducted in a controlled and limited way. A porous frontal defense would have enabled Warsaw Pact forces to advance in many areas along a broad front. A geographically widespread use of nuclear weapons might have been necessary to halt them. The layer cake, by contrast, promised to help enable NATO to contain the enemy's advance along narrow salients where Warsaw Pact forces had concentrated and broken through. These salients potentially could be closed by limited NATO nuclear employments of the sort that would not inflict massive destruction around the larger German countryside. This situation, of course, would not ensure against a subsequent widespread escalation if Soviet forces responded in kind, but it did promise to give NATO at least the option of trying to fire a militarily meaningful shot across the bow without automatically touching off a full-scale, theater-wide nuclear war. For all these reasons, the layer cake underwrote not only the direct defense provision of MC 14/3, but also the concept of deliberate escalation.

[7]See Kugler, 1991.

Military Reasons for Linear Defense

Purely military reasons also drove NATO's planners toward a linear defense and layer cake array. Figuring heavily in NATO's calculus were the disadvantageous physical dimensions of the inter-German border. The origins of the problem began at the end of World War II. As of May 1945, when the war in Europe ended, allied forces had advanced well eastward of what later became the inter-German border. The British 21st Army Group and the U.S. 9th and 1st Armies stood only 60 miles from Berlin. To the south, the U.S. 3d and 7th Armies occupied a line from Chenmitz, 40 miles south of Dresden, through Pilsen in Czechoslovakia, to Linz, Austria. This situation gave allied forces control over the Elbe River line, the densely forested Thuringerwald, and the mountainous regions of western Czechoslovakia, highly defensible terrain.[8]

In honoring earlier agreements with the Soviets, allied forces withdrew to what became the inter-German border, thereby sacrificing this advantageous terrain. This withdrawal left them holding a lengthy, concave defense line with exterior (long) lines of communication against a Soviet Army that held the advantage of a convex line and interior (short) lines of communication. Additionally, allied forces now had little strategic depth in which to maneuver; the Paderborn Plain, the industrial Ruhr, and the Rhine River lay only 200–300 km away. In other words, the Western allies now found themselves facing the worst possible geographic situation for a heavily outnumbered force attempting to contain a larger, offensively capable attacker.

Compounding this problem was the evolutionary course that Soviet/Warsaw Pact forces pursued during the 1950s and beyond, which strengthened their capacity to exploit NATO's disadvantageous situation. During this period, the Soviets embarked on an effort to upgrade the 20-division Group of Soviet Forces, Germany (GSFG) in East Germany. In 1968, five Soviet divisions were permanently deployed in Czechoslovakia in the aftermath of the invasion; this force was entitled the Central Group of Forces. This step, coupled with the two Soviet divisions deployed in Poland as the Northern Group of Forces, gave the Soviets a total forward-deployed posture of 27 divisions. Added on top came the effort, undertaken in the mid-1950s, to establish East European armies and to integrate them into the Warsaw Pact. In total, some 31 divisions were created in Poland, Czechoslovakia, and East Germany. These forces were gradually modernized and acquired an offensive capability. Finally, the Soviets began strengthening their roughly 33 divisions in the three immediately adjacent western military districts: the Baltic, Byelorussian, and Carpathian MDs. In time, these forces were incorporated into Soviet/Warsaw Pact war plans for operations in Central Europe.[9]

Warsaw Pact force expansion efforts levelled off by the late 1960s as total force levels opposite NATO's Central Region stayed roughly constant from then on. However, a major modernization program was launched to carry forth the pace of steady improvements and broaden Soviet military options. Before that, Soviet/Warsaw Pact forces primarily had been configured for nuclear war. While they were equipped with large numbers of tanks, they lacked the other elements of the combined arms team that were needed for conventional combat. This situation soon changed as Soviet/Warsaw Pact forces began acquiring the self-propelled artillery; infantry fighting vehicles; larger infantry formations; attack helicopters; tactical combat aircraft; command, control, communications, and intelligence (C3I) systems;

[8]See Eisenhower, 1986.

[9]See Wolfe, 1970.

and logistic support forces needed on the modern battlefield. This buildup was launched in the mid-1960s, at about the same time that the Soviet nuclear buildup began gathering steam. By the mid-1970s, its effects were being decisively felt in Central Europe.

The most visible effect was to erase the qualitative edge in weapons and forces that NATO had traditionally counted on to offset its numerical inferiority. Previously, NATO's tanks, artillery, helicopters, and other weapons had been markedly better than their Soviet counterparts. By 1980, Warsaw Pact weapons matched the West's best. Previously, also, NATO's ground divisions had been larger and better armed than the Warsaw Pact's. This advantage too had largely disappeared: Soviet divisions, while still having somewhat fewer personnel, now possessed comparable weapons inventories in virtually all categories. Before this buildup, despite being outnumbered by 2:1 in DEs, NATO trailed by a less decisive margin of 1.5:1 in total weapons and firepower (as measured in WEI/WUV terms). By 1980, as Fig. 2 shows, the WEI/WUV disparity itself had grown to nearly 2:1.[10] In other areas, such as tactical air power, command and control, training, and leadership, NATO still enjoyed important qualitative advantages. But even with these, the 2:1 disparity in weapons was worrisome.

In addition to being outgunned, NATO also found itself facing an adversary force posture that was increasingly being brought into alignment with the requirements of an

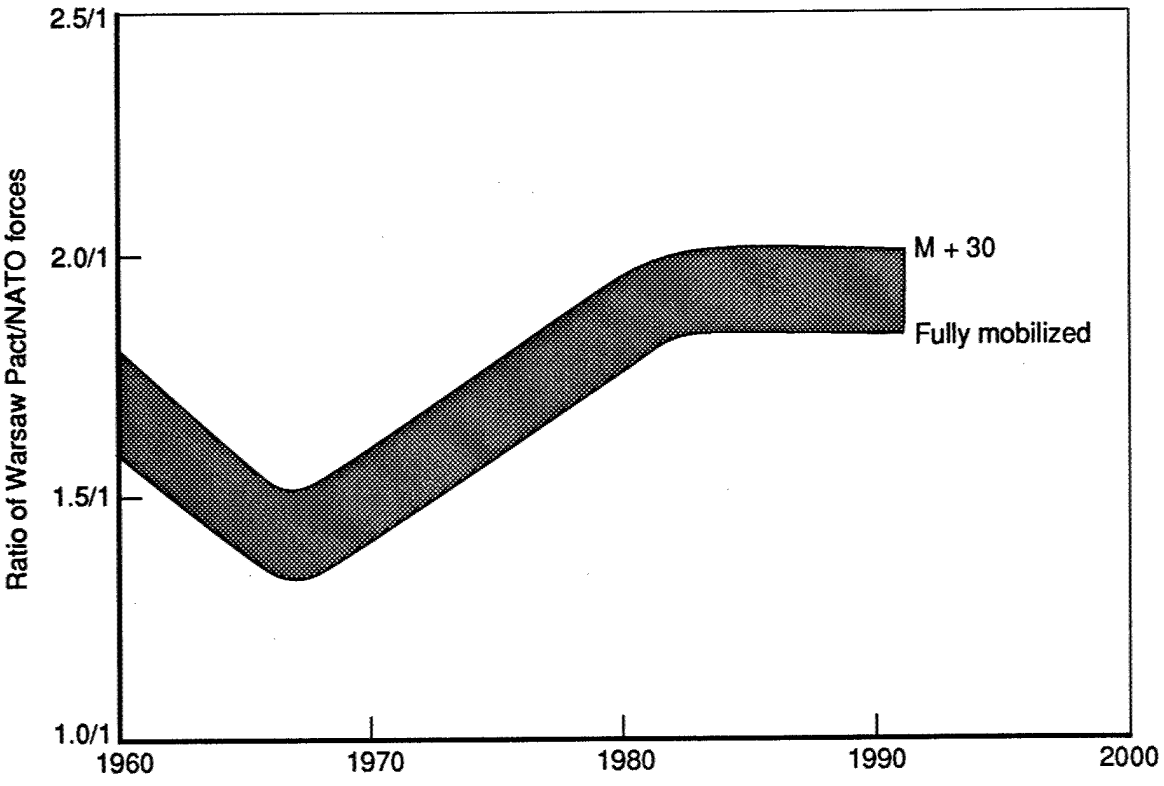

Fig. 2—Historical Trends in WEI/WUV Scores

[10]For a further analysis, see Karber and Whitley, "The Operational Realm," in Golden et al., 1989.

offensive military strategy. In the early 1960s, Warsaw Pact forces were essentially one-dimensional: Oriented to nuclear war, they were not well designed for carrying out the complex maneuvers associated with conventional campaign plans. By 1980, this had changed. Warsaw Pact forces were now able to perform both highly concentrated breakthrough operations against prepared frontal defenses and swift-moving envelopments in the rear areas. As a result, they could launch and carry out a purely conventional operation to completion, one aimed at gaining a swift and decisive victory.

This military buildup left NATO's planners facing the problem of defending against several different kinds of Soviet/Warsaw Pact attacks in a crisis. All of these attack options had to be taken into account in the preparation of NATO's forces and employment plans, thereby inhibiting NATO from optimizing against any one of them. The most immediate risk was that of a surprise, "bolt out of the blue" attack across the inter-German border after only a short period of warning. Since Soviet forces in East Germany were maintained at high peacetime readiness, NATO planners feared that such an attack might be launched within as little time as 4–5 days after a Warsaw Pact mobilization. The 20-division GSFG was viewed as the main force for a short warning attack, with the 2d Guards Tank Army, the 3d Shock Army, and the 8th Guards Army, all based near the border, serving as the vanguard. Additional Soviet, East German, and Czech forces seemed capable of bringing the total attacking posture up to about 40 divisions, or two "fronts."[11]

These forces themselves were not large enough to overpower the Alliance's defenses, but if NATO's forces failed to mount stiff resistance at the border, they threatened to march a short distance across the inter-German border and seize nearby cities, thereby inflicting an embarrassing setback on NATO and the West German government. Concern about a limited attack of this sort played a role in NATO's adoption of MC 14/3. Even more worrisome, these Warsaw Pact forces also threatened to advance more deeply and to disrupt NATO's mobilization efforts. In the immediate aftermath, follow-on Warsaw Pact reinforcements might then be able to advance successfully against a disorganized NATO defense.

Although NATO's military commanders came to fear this kind of "short warning" attack, the Soviets also had the option of taking a longer period to mobilize and reinforce to build up a much larger posture. Drawing on the East European forces alone, the Soviets were capable of deploying some 60 Warsaw Pact divisions, enough units to permit a three-front attack capable of exerting offensive pressure along NATO's entire Central Region front. Reinforcement by the Soviet divisions from the western MDs promised to bring the total up to about 90 divisions. This posture would have provided a five-front offensive, with two fronts used as a second strategic echelon capable of exploiting an advance deep into NATO territory well beyond the Rhine River. Additional Soviet units were available from more distant military districts, potentially bringing the Warsaw Pact posture up to 120 divisions.[12]

These larger options would have required additional time to mount, well beyond the 4–5 days needed for a short warning attack. Many of the extra Soviet/Warsaw Pact divisions were maintained in cadre status in peacetime, with only partial manning by active duty personnel. These units would have required a period of time for mobilization and refresher training. Also, the physical task of moving them forward from the rear areas promised to consume about two weeks. Over the years, Western planners studied the issue of exactly how much time would be required for the entire effort. Estimates ranged from as little as

[11]See John Erickson et al., 1986.

[12]See Secretary of Defense, Annual Reports to the Congress (DoD Posture Statements), Washington, D.C., 1975–1988.

two weeks for a 90-division assault up to a month or even longer. Regardless, two considerations stood out. First, the Soviets had a major incentive for attacking quickly, before NATO had an opportunity to mobilize its impressive military, industrial, and economic assets. Second, despite its staggered readiness profile, the Soviet/Warsaw Pact posture was sufficiently ready to attack before NATO could create entirely new forces out of wholecloth. NATO had to plan to fight with its forces in being, some of which would have to be transported from the United States.

Complicating NATO's efforts to defend against any of these large enemy attacks was the terrain along the inter-German border. The border area by no means has been indefensible, but it offered the Soviets several avenues for high-speed armored spearheads. Moreover, these attack corridors are spread up and down the inter-German border. The most dangerous terrain, from NATO's standpoint, was in the NORTHAG area, where NATO's defenses traditionally have been weakest. The North German Plain, south of Hamburg, provides a direct line of march over open, rolling terrain to the Paderborn Plain and the Ruhr. Further south, the autobahn along the Hannover-Braunschweig axis and the Gottingen Gap through the Harz mountains also point directly at the Ruhr and the Rhine River.

The CENTAG terrain is less open, but it still has several attack corridors, particularly the Hessian Gap; the Meiningen Gap; and the Hof, Cheb, Coburg, and Highway 14 approaches. Only the Bavarian forests to the south, guarded by the thinly spread II German Corps, denied opportunity for concentrating at least 2–3 heavy enemy divisions along a single narrow axis. This area aside, NATO's forward corps sectors have contained at least one, sometimes more, attack corridors that needed to be guarded in strength. The geographically widespread distribution of these corridors complicated NATO's defense plans by compelling the Alliance to pay attention to virtually the entire Central Region front. This requirement, in turn, exerted a powerful constraint on NATO's ability to thin out its defenses in some areas to concentrate powerful forces in a few sectors. Figure 3 illustrates the locations of these avenues of attack.

In designing employment plans for defending the Central Region against the multiplicity of enemy attack options over this terrain, NATO's planners in the 1950s initially were not committed to the concept of a linear defense. This was true even in West Germany, where Bundeswehr officers tended to favor a mobile defense of the sort that the Wehrmacht had conducted on the Russian front in World War II. But once enough NATO forces had been deployed to make a linear defense feasible, a military consensus built in favor of the layer cake. The reasons deserve explication here.[13]

Basing arrangements were driven by such practical considerations as available facilities, demography, national predilections, and the accidents of history. NATO's combat and logistic support forces wound up being distributed in a generally linear fashion on a north-south axis up and down the West German countryside. This pattern left the Alliance well situated for forming a linear defense and not geographically well disposed for any employment doctrine that required the prompt massing of NATO's forces.

Another factor was that a quickly constitutable linear defense offered NATO military advantages for contending with a limited "short warning" Soviet attack, especially one aimed at seizing exposed West German territory. A doctrine that called on NATO's forces to move directly forward to the inter-German border, in a corps-abreast formation, placed NATO in position for blocking a standing-start GSFG attack launched westward from Soviet bases in

[13]See Kugler, 1991.

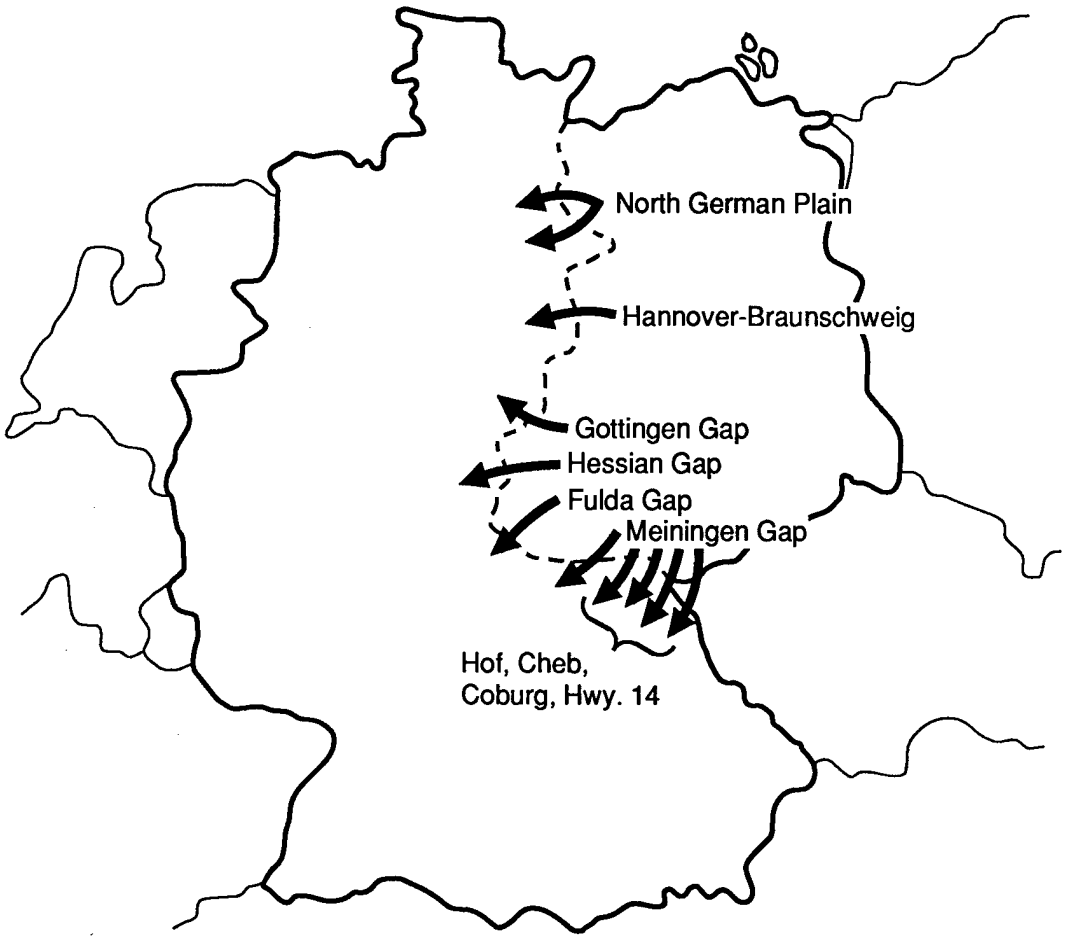

Fig. 3—Avenues of Advance Along the Inter-German Border

East Germany. See Fig. 4. As a result, the GSFG was denied the capacity to conduct a broad-front attack aimed at gaining easy, limited advances along the border.[14]

Linear defense offered a way to contain the larger Soviet attack that might be launched later (90 divisions at M+15–30). For that scenario, the layer cake provided an employment doctrine that enabled NATO's outnumbered forces to support each other in an organized fashion. The alternative to the layer cake was a mobile defense in which large armored forces, initially withheld as operational reserves, would function as NATO's primary vehicle for stopping an enemy attack. Only a thin front line would be formed, with just enough forces to screen, delay, and channel enemy forces toward killing zones in the rear. There, counterattacking NATO forces would employ "hammer and anvil" tactics, meeting engagements, flanking attacks, and other maneuvers to destroy enemy formations.

[14]See Erickson et al., 1986, p. 38, for Soviet deployments. For NATO deployments, see Chailand and Rageau, 1985, p. 214.

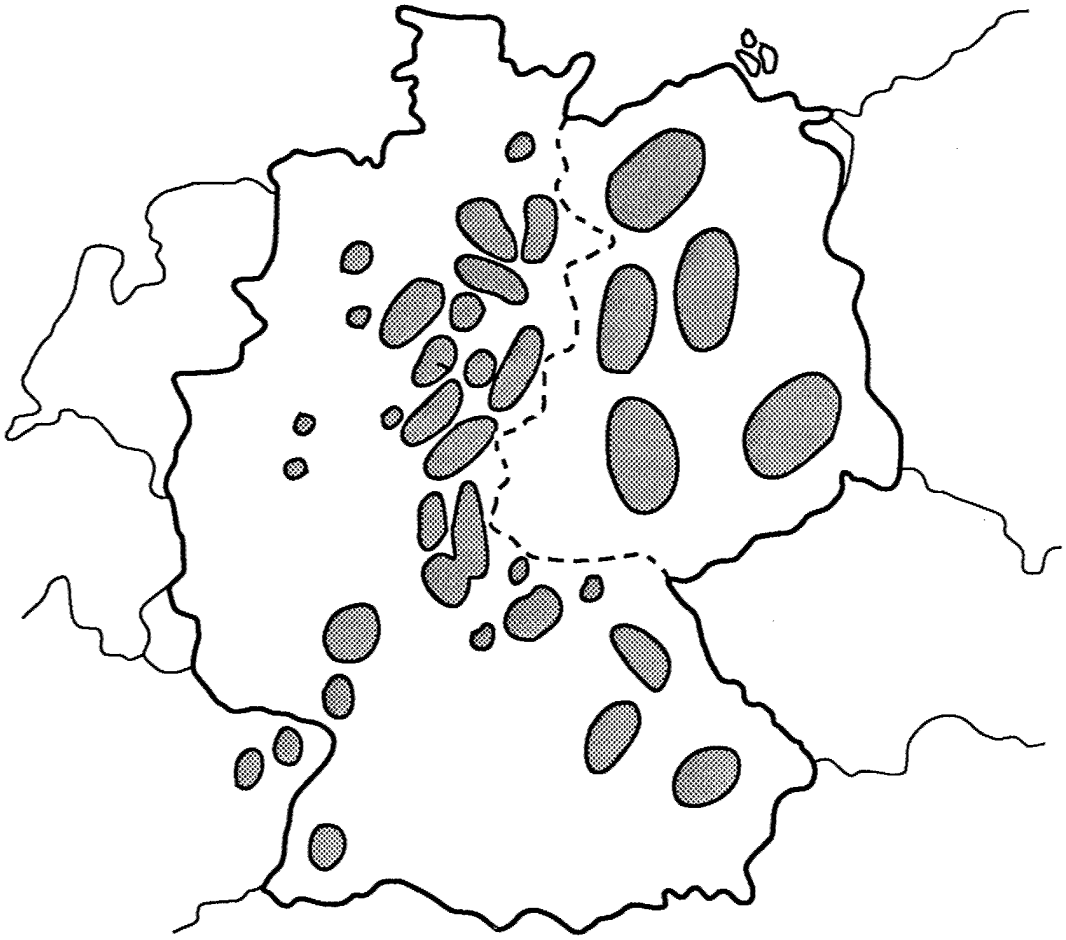

Fig. 4—Deployment Patterns of Soviet and NATO Forces

Mobile defense's advantage lay in its promise of using possible NATO advantages in training, leadership, and tactics to offset the large numerical disparity facing AFCENT. Above all, it promised to make NATO's posture less brittle and subject to sudden, complete collapse. Its disadvantage lay in its willingness to sacrifice ground to set up the enemy for counterstrokes, territory that NATO politically could not afford to cede. Additionally, mobile defense presumed that NATO's edge in these areas would be sufficient to offset a numerical disadvantage of as large as 2:1. To the extent this edge proved illusory, or merely insufficient, a mobile defense ran the risk of magnifying, rather than reducing, NATO's military problems.[15]

In particular, a mobile defense threatened to transform the Central Region battlefield into the kind of maneuver war that the side with the largest forces normally seeks because he is able to bring his strengths fully to bear. In open combat where terrain does not constrain maneuvers, the side with larger forces normally enjoys broader options than the

[15]For a more detailed appraisal of mobile defense, see Kugler, 1989.

smaller contender. Above all, larger forces enable that side to seize the initiative, encircle the opponent, break apart his forces, and defeat them in detail. For this reason, mobile defense, as a theoretical concept, was not seen as a cure-all. In Central Europe, it offered an uncertain prescription for NATO's defense ills, one that the Alliance, for military as well as political reasons, chose not to swallow.

Linear defense was anchored on a different but also plausible theory for fighting outnumbered and winning, one that proved more palatable to the Alliance. A frontal wall provided NATO a vehicle for closing off the terrain to the Soviets. By denying the Soviets depth to maneuver and spread out, it promised to jam up Soviet attack formations, compelling them to advance in successive echelons. This had the beneficial effect of reducing engaged force ratios to more manageable portions, easing NATO's defense task.

A frontal posture also promised to give NATO the classical advantages that accrue to the defender and help offset the attacker's edge in possessing the initiative. A frontal posture enables the defender to fight on preselected terrain. If enough time is available for engineering construction, it also permits defensive operations from prepared positions. Further, it enables a defender to fight with its flanks secure and under controlled conditions. As a result, every unit is able to focus its weapons and firepower in one direction, from which the enemy attack is most likely to come.

A frontal posture also provides the defender a free zone in the rear areas, where command and control centers can be established, maintenance performed, supplies sent forward, and other logistic support functions provided. Also, it gives a clear demarcation line for employing tactical air power in close support roles. A frontal posture thus enables a defender to employ all his assets—combat, combat support, and combat service support—to full advantage. It enables the defender to generate a large volume of highly controlled, accurately delivered, and sustainable firepower, which can enable an outnumbered defender to attain the advantageous exchange rates that are needed to burn away the attacker's superior numbers.[16]

Operationally, a frontal defense promises to transform combat into a set-piece war of attrition. Combat might be conducted with frenetic intensity, but the attacker normally does not advance into the defender's territory quickly. Victory ultimately is determined by which side outlasts the other through a process of grinding the victim down into submission. A frontal defense is no cure-all either. But when the defender's forces are large enough to keep the balance within manageable limits, a frontal posture can be a sound prescription for effective defense, especially in support of a military strategy that cannot afford to yield much ground.

For this reason, defending armies often have attempted to establish a strong frontal wall, physically akin to NATO's layer cake array, when confronted with a dangerous attacker. Mobile defense has been chosen when the front was too long and the defender's forces too small to form an adequately dense front line, or when the defender had a clear qualitative edge in maneuver combat. But when a frontal defense was physically possible, it often was the preferred employment doctrine for meeting the enemy's initial assault.

NATO formally embraced the frontal defense philosophy when it selected the layer cake concept in the 1960s. From that point forward, its defense programs became focused on the task of deriving maximum firepower—immediate and sustained—from a posture whose size was fixed. This does not imply that the Alliance was oblivious to plans to create more forces

[16]For a appraisal of how NATO's frontal defense doctrine affected procurement programs, see Richard L. Kugler, "On Land Warfare: NATO's Quest for Vision, 1962–1986," in Koldoziej and Morgan, 1987.

and to enlarge its defenses in other ways. The U.S. POMCUS program and the expansion of the West German Territorial Army are programs aimed at increasing NATO's forces and providing greater depth in the rear areas, but the bulk of NATO's improvement programs were directed toward maximizing NATO's firepower. Its force goals of the late 1980s are no exception.

Drawbacks of NATO's Linear Defense

Notwithstanding its advantages, NATO adopted the layer cake with some misgivings. One fear was that it might initially be manned by insufficient forces to block a short-warning attack. If NATO's decision to mobilize lagged behind the Soviet decision by more than a day or two, only 10–20 divisions would have been available when Soviet forces crossed the border. These forces might not have been enough to block the attack.

A more prominent concern was that the layer cake's lack of operational reserves would prove to be NATO's Achilles heel against a fully mobilized attack. Without sizable reserves, even a strong frontal wall can be destroyed by a larger attacking force through sustained pounding. More likely, it can be breached at selected points by concentrated attacks that enable enemy spearheads to advance into the rear areas and then to unhinge the defense through maneuver operations. The defender requires reserves to block, contain, and otherwise stymie these breakthrough operations before they get beyond control.

Ever since the layer cake was formed, a perennial issue has been whether NATO would have enough operational reserves for this purpose, especially if a war were to begin before the U.S. buildup effort is complete. The central problem was that the layer cake consumed 32 NATO divisions, with 24 divisions assigned to frontline roles and another eight divisions used as corps reserves. The guiding concept was to ensure that all eight of NATO forward corps sectors could hold off a major attack for several days, long enough for help to arrive from elsewhere. As Table 2 shows, this allocation left NATO with few forces that could be withheld in the rear areas.[17]

The lack of reserves seemed especially serious during the early stages of mobilization and buildup, although the deployment of U.S. reinforcements promised to reduce this deficiency in later stages. Once the U.S. buildup had been completed, NATO's posture would provide the 50–60 divisions that the Alliance's force goals traditionally have demanded. It

Table 2

NATO Ground Force Levels in AFCENT (DEFs, 1989)

	Forward Corps	Operational Reserves	Total
M+7	32	5	37
M+15	32	10	42
M+30	32	13	45
M+45	32	18	50
M+60	32	22	54
M+75	32	25	57

[17]See Steinbrunner and Sigal, 1983.

also would reduce the Warsaw Pact to NATO force ratio nearly to the 1.5/1 standard that many analyses have postulated as necessary for a confident defense against a 90-division attack. The U.S. buildup, however, appeared likely to take two or three months to complete. The fear that afflicted NATO during the Cold War was that a Soviet/Warsaw Pact attack would have been launched well before then.

Appraisal of NATO's Conventional Defenses Prospects

What was the military significance of this shortfall? In the 1950s, NATO's governments generally conceded that the Alliance's conventional defenses in Central Europe could have been overrun quickly. This pessimism gave way to a more optimistic appraisal in the late 1960s, especially the view of U.S. Secretary of Defense Robert McNamara, who insisted that conventional defense was a viable goal. Nevertheless, even McNamara called for further NATO force improvements, including larger reserves. In succeeding years, official appraisals took a more negative turn. By the mid-1980s, SACEUR General Bernard Rogers was warning that largely because of NATO's lack of reserves and related shortfalls in ammunition stocks and other areas, NATO might be compelled to resort to nuclear weapons within a week or two after D-Day.[18]

The U.S. Department of Defense broadly echoed this concern, if in less stark terms. Certain American academics reached more optimistic conclusions, and their voices were heard with greater frequency in the Cold War's waning years. Their views aside, however, the dominant Western consensus for some years has been that during the Cold War, the Central Region balance was unfavorable to NATO.[19]

RAND studies have supported this conclusion. Thomson (1988) stated, "the imbalance of forces in Central Europe favors the Warsaw Pact today. NATO would probably lose most war scenarios and be decisively defeated in many." Davis (1988a) dwelt more heavily on the uncertainties in analyzing the balance. But he called for increased operational reserves and other measures to remedy acknowledged weaknesses in NATO's posture.

These conclusions have been influenced by dynamic computer modeling at RAND in recent years. For example, Fig. 5 displays the results of an analysis employing the RAND Strategy Assessment System (RSAS), a fully automated computer simulation tool that allows for sensitivity analysis of multiple scenarios and battlefield situations.[20] This model is particularly useful in examining how local pressures in breakthrough sectors could have compelled NATO to retreat along the AFCENT front to shore up exposed flanks. In the canonical scenario, NATO's forces were driven back deep into West Germany within a month of D-Day.

Figure 6 displays the results of a RAND/Arroyo Center study that used the IDAHEX wargaming system, which allows for detailed evaluation of single scenarios. It is especially useful in examining how the Soviets could have exploited the Central Region terrain and battlefield dynamics to gain breakthroughs against a defender that is stretched thin. The result was a catastrophic breakthrough in the NORTHAG sector after about 2–3 weeks, producing a cascading sequence of events leading to NATO's unravelling and operational defeat

[18]See Secretary of Defense Casper Weinberger's Annual Reports to Congress for FY1983–1988. See also Joint Staff, 1989.

[19]For a more optimistic appraisal, see the works of Posen, 1984; Epstein, 1987, 1990; and Mearsheimer, 1983. A good example can be found in *International Security*, Vol. 12, No. 4, Spring 1988.

[20]The RSAS includes interactive wargaming, along with use of the CAMPAIGN computer model, to explore adaptive strategies and the role of maneuver operations in linear or other types of defense efforts. Figure 5 is one illustration of the types of military conflicts that this system can simulate, including those involving substantial maneuvers.

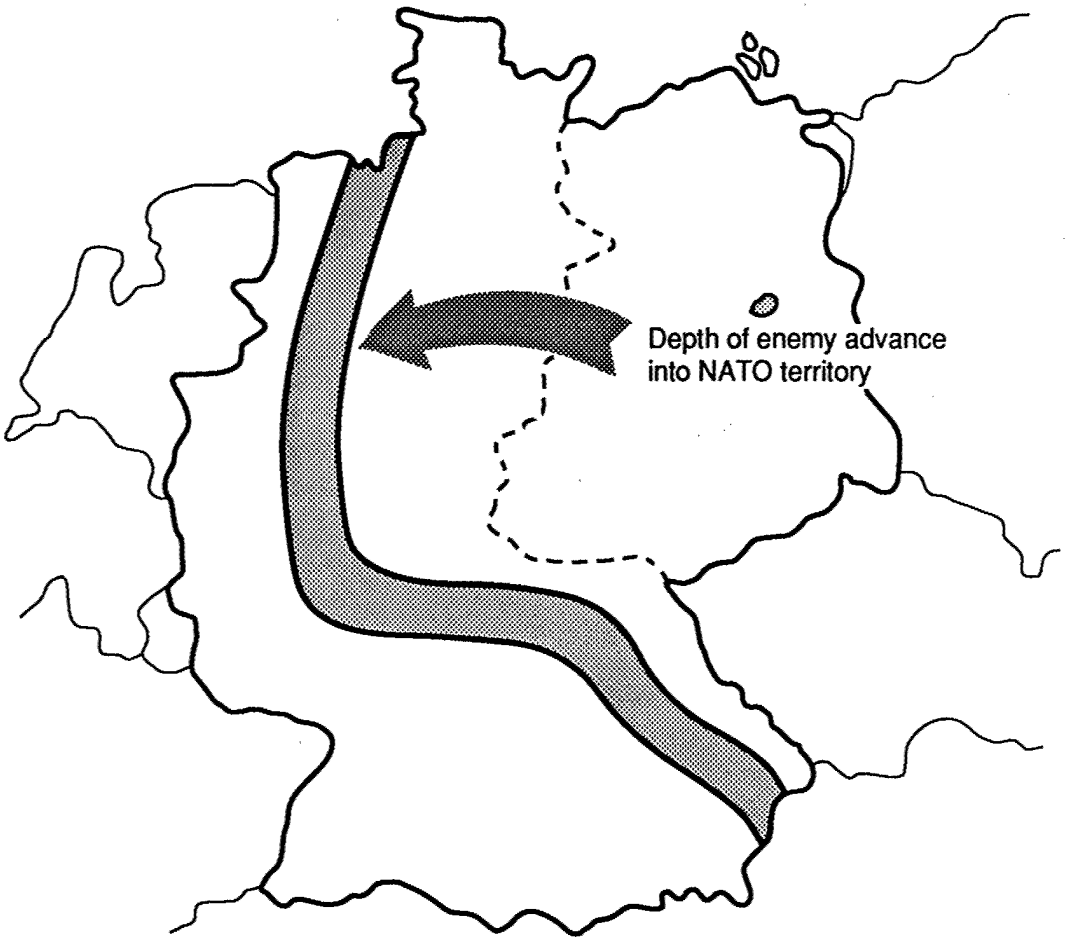

Fig. 5—Illustrative Campaign Results

there. As in the CAMPAIGN study, NATO did much better in CENTAG, where NATO's forces are stronger and the terrain is more favorable. But the defeat in NORTHAG dominated the outcome.

These two studies reached similarly negative conclusions with two different models, each examining NATO's situation from complementary angles of vision. Their results, alone or together, by no means imply that NATO has been foredoomed to failure. Both studies were based on scenarios that reflect consciously conservative assumptions: e.g., a quick Warsaw Pact buildup followed by an effectively prosecuted warplan in which both Soviet and East European forces perform at high levels of effectiveness.

Also, computer simulations typically ignore (or make neutralizing assumptions) about a whole set of largely exogenous variables that can have an important influence on the outcome. Included are leadership, morale, small-unit training, and tactics. These factors introduce enormous uncertainty into the evaluation. To the extent that they would have operated in NATO's favor, the West's prospects would have improved. In this context, history shows many cases in which an outnumbered defender not only halted but also decisively defeated a

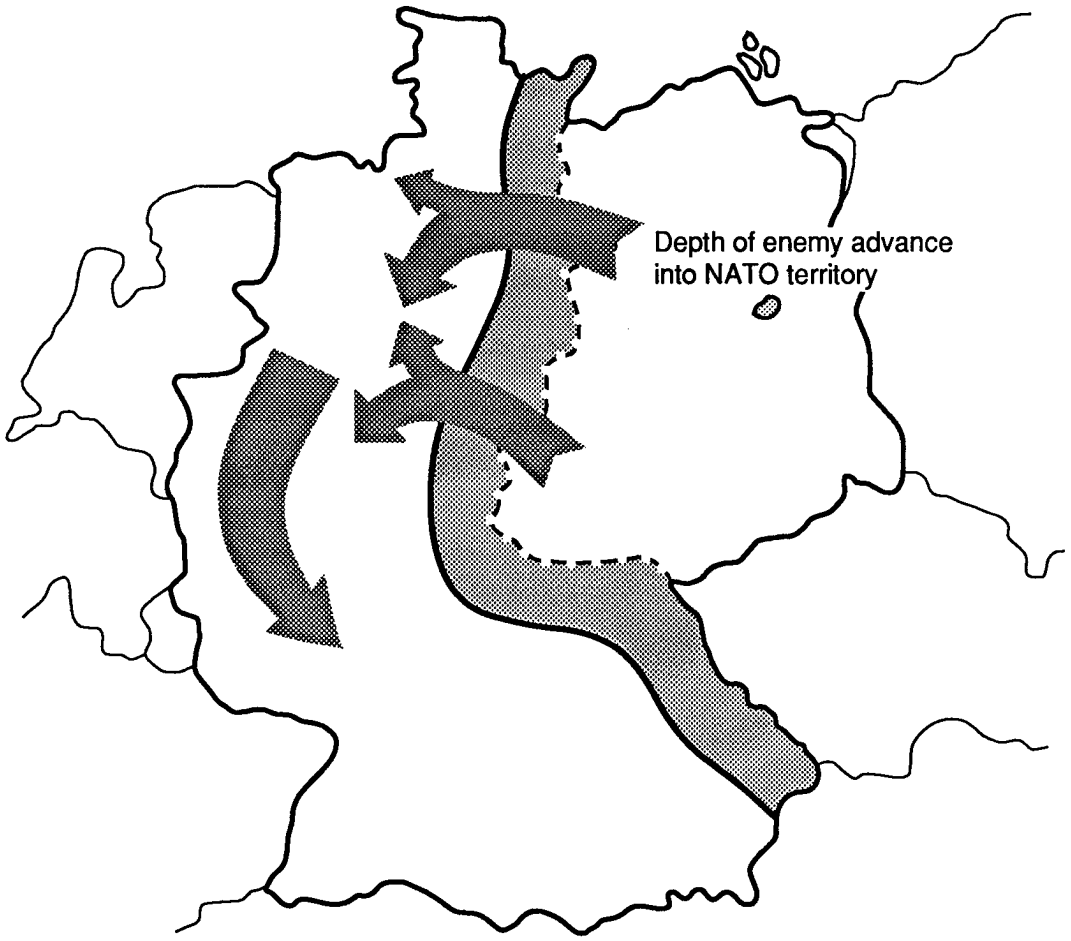

Fig. 6—Illustrative IDAHEX Results

larger aggressor. Had war broken out in Central Europe, NATO's forces in all likelihood would not have been overwhelmed quickly and easily. Nevertheless, by most analytical measures NATO's posture has not met the tests of sufficiency that commonly are employed to gauge a high-confidence defense. NATO's vulnerabilities, over the years, have created risks that have seemed uncomfortably high to most Western observers.

For this reason, the Central Region conventional force balance has not been "militarily stable" as this term is employed here. But neither has it been wholly unstable. NATO's conventional posture, despite its weaknesses, has been strong enough to play an imperfect but still contributing role in achieving the important objective of deterrence. Although this posture has left NATO with a sense of insecurity, it has been strong enough to deny the Soviets high confidence in their own prospects of prevailing in a war. It thus has removed obvious temptations to aggression and has left the Soviets facing incentives that only the risk-oriented could find enticing.

Despite its weaknesses, NATO's linear defense array (and the forces that implemented it) has been strong enough to achieve at least MC 14/3's minimum standard for forward

conventional defense. It enabled the Alliance to assert political control over the inter-German border in peacetime and in crises. Despite its vulnerabilities, it provided NATO at least an initial period of direct defense. It gave NATO the important window of time on which the decision to cross the nuclear threshold had rested. The layer cake thus helped give life to MC 14/3's other, equally important concepts of deliberate escalation, extended deterrence, and general nuclear response.

By constructing at least an initial defensive capacity of a few days or weeks of staying power, NATO's linear defense has also helped the Alliance pursue other important security objectives. Internally, it has helped meet the U.S. demand for a conventional foundation to flexible response and thereby has helped ensure the continued U.S. commitment to extended deterrence. Meanwhile, it has helped assuage West German concerns about the Alliance's commitment to forward defense. It encouraged coalition planning by bringing several nations together in a common enterprise, helping preserve NATO's internal unity during years in which the Alliance was vulnerable not only to external manipulation but also to its own internal strains.

In a similar vein, NATO's linear defense plan helped to politically stabilize European security affairs. Not entirely by accident, the Soviet government stopped fomenting periodic tensions and crises in Central Europe once a coherent NATO posture and layer cake employment plan took shape. This objective was achieved with a force posture that was not provocative and therefore helped to support NATO's arms control goals in Europe. Finally, the layer cake proved to be affordable. Throughout the final years of the Cold War, NATO was beset with growing strains over burdensharing issues, but these strains were never powerful enough to fracture the Alliance.

Where NATO's linear defense has fallen down is in its capacity to achieve MC 14/3's maximum goal of a fully stalwart conventional defense, one that could sustain itself for a lengthy period and ultimately stalemate even a massive Soviet invasion. The Alliance's members collectively never judged the goal of a stalwart defense to be sufficiently important to justify the additional funds and commitments that would have been necessary to achieve it. The idea was primarily an American conception. The other allies supported it less warmly; in earlier years, some even regarded it to be questionable. In purely subjective terms, NATO's failure in this area thus was not a crippling one.

At the same time, in objectively measured security conditions, the Alliance paid for its failure to field stronger conventional defenses. To some degree, deterrence was weakened. In this regard, the Alliance was fortunate that a severe crisis in Europe never coincided with the presence of a risk-oriented Soviet leadership. In ways that historians will never be able to measure precisely, NATO's military deficiencies were probably offset by political luck. Likewise, the Alliance for many years lived with a greater reliance on nuclear escalation than many observers deemed prudent. This too was a cost, but one that is hard to measure in precise terms. What can be measured more precisely is the political cost to NATO's internal unity that this nuclear reliance inflicted on the Alliance and on the transatlantic relationship. Although NATO endured the Cold War with its cohesion generally intact, the journey was made considerably more stressful by this problem. In this area, also, NATO is probably fortunate that it survived in the ways it has.

FUTURE SECURITY CONDITIONS IN EUROPE

If this appraisal characterizes the Central Region military balance over the last two decades, how are security conditions in Europe likely to change now that the Cold War has ended? How will NATO's military situation be affected and what implications will be posed for NATO's employment doctrine?[21]

Figure 7 provides a summary portrayal of where the European security system seems to be headed over the next 5–10 years or so. In past years the security system was structurally bipolar, with two cohesive military blocs confronting each other at the inter-German border. This old order is rapidly giving way to what might best be called a depolarized security system. On the western side is a still-existing NATO, with united Germany as a member. On the eastern side is the Soviet Union, with its military forces withdrawn behind the USSR's

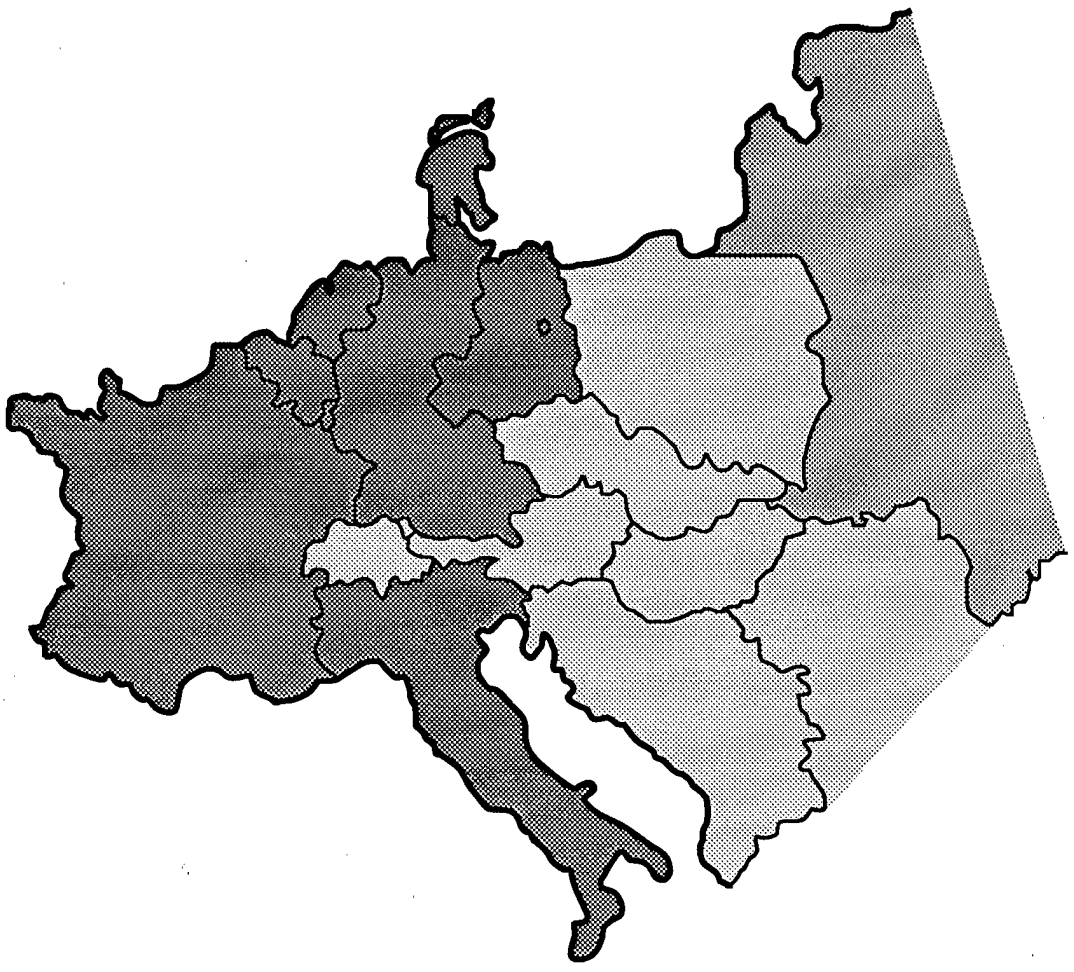

Fig. 7—A Future Depolarized European Security System

[21]For a useful analysis, see Davis, 1990.

border and at best tenuous security ties to its eastern neighbors. In the middle is a collection of largely neutral East European nations where the Warsaw Pact once stood.[22]

The virtual death of the Warsaw Pact does not mean the severing of all Soviet-East European security connections. For example, the USSR might well preserve security ties to the Balkan nations, including Bulgaria and possibly Romania. Future Soviet relations with Poland, Czechoslovakia, and Hungary are unclear. Perhaps these nations will maintain some security ties with the USSR under the rubric of existing mutual defense treaties. But this caveat aside, these three nations will be acting more independently in their foreign policies. They are portrayed here as pursuing closer diplomatic and economic relations with the West, and possibly security ties as well. The Warsaw Pact no longer will be an integrated military alliance under Soviet control. In this depolarized system, Eastern Europe functions as a military power vacuum and buffer zone between NATO and the USSR.

Figure 7 is intended to be a heuristic device, not a fixed predictive blueprint for the future. Nor is it intended to represent a final stage in Europe's evolution; further important changes, possibly toward a collective security era, might lie ahead. Also, at the time of this writing, Europe has not yet fully become depolarized. Among other things, sizable Soviet forces remain in Eastern Europe that must leave the forward areas before this system can come to pass. A major, perhaps difficult, transition lies ahead. Nevertheless, the changes at work over the past couple of years seem to be pointing strongly in the direction of a depolarized system, and it is noteworthy how far events already have gone in this direction. The West intends a unified Germany to remain in NATO's integrated military structure, the East European nations have shaken off the shackles of Warsaw Pact controls, and the Soviet Union has committed itself to troop withdrawals. To the extent that Europe's future can be foreseen with any confidence, a depolarized system provides a satisfactory basis for thinking about the West's security policy in the post–Cold War era of a few years from now.

What is striking about the depolarized system is the dramatically different security environment it will create for NATO's diplomacy and defense policy. During the Cold War, the West's main goal was to manage the bipolar confrontation at the inter-German border. The Alliance's security policy primarily was one of containment and deterrence, supplemented by efforts to reach accommodation through arms control negotiations and political accords. A depolarized system, with its different features, will enable (indeed, demand) that NATO's security policy acquire a broader conceptual and geographical emphasis. Rather than concentrating on the inter-German border, NATO will need to expand its vision to include all of Europe, and it will need to interweave its diplomatic and economic endeavors with its defense efforts. Similarly, NATO's policy aims will need to expand beyond the narrow confines of containment and deterrence to address the more nebulous but equally important task of erecting a stable balance of power and security architecture in Europe.

The concept of a "balance of power" policy aimed at building a stable security architecture implies a careful blending of coalition strength with a recognition of the legitimate interests of the nations east of NATO's borders. A core feature of this security policy would be a transformed stance toward the Soviet Union. The continued existence of partly conflicting goals would still be recognized, but the USSR would also be treated as a nation with valid interests of its own and with which growing forms of cooperation are feasible. At a minimum, the West would aspire to create a mixed, pragmatic relationship with the Soviet Union that entailed a combination of still-remaining rivalry and enhanced cooperative partnership.

[22]See Kugler, 1991.

This kind of relationship now appears possible. Throughout the Cold War, the Soviet Union was largely viewed as an implacable, ideologically motivated, and expansionist adversary. The Gorbachev reforms have altered the USSR's policy in the direction of a defensive strategy, one motivated by the Soviet Union's own internal economic weaknesses and political difficulties. Added on top has come the defection of the East European allies, depriving the Soviet Union not only of its buffer but also the forward-based springboard that the Soviet government could have used to invade the West. As a result, NATO's defense planning no longer will need to be preoccupied with a short-warning attack directly onto German territory once Soviet withdrawal is complete. Equally important, the USSR's political agenda in Europe appears less threatening and more oriented toward building stable relations on terms respectful of the West's interests. To the extent that the Soviet Union continues pursuing a more benign external policy while democratizing internally, cooperative relations with the USSR will become increasingly possible.

Provided these trends are neither halted nor reversed, the Europe of the future seems likely to be more stable, less tense than that of the past. At the same time, security issues are unlikely to disappear from Europe's agenda, especially in some wholesale way. In this regard, it is useful to recall that since Napoleon's time, Europe's security affairs have gone through a pattern of regular cycles in which old problems disappeared only to give way to new ones. At each juncture, observers often concluded that enduring tranquility was at hand, only to be disappointed because they failed to recognize the fault lines ahead.

The Concert of Europe maintained peace from 1815 to 1854 by joining the leading conservative powers in a pact to resist revolutionary change. Although often characterized as Europe's golden age of Metternichian parlor diplomacy, it experienced strains aplenty. Unable to contain the forces unleashed by industrialization and emerging nationalism within individual European nations, it ultimately collapsed when it could not accommodate German and Italian unification. Replacing it came the Balance of Power system, which sought to maintain stability through a mechanical process of counterbalancing alliances and military power blocs. Wracked by nationalistic rivalries and competitive arms races, it exploded suddenly into World War I in 1914. Following the war's end in 1919 came a 20-year pursuit of a collective security structure in the League of Nations, which failed to contain totalitarianism on both ends of the political spectrum and ultimately ushered in World War II. The Cold War from 1947 to 1989 was not a historical anomaly. Rather, it reflected an onward march of European historical patterns that, if the past is prologue, is likely to repeat itself in the coming years.[23]

Although its specific features are anything but clear, the depolarized system seems likely to present an array of political problems of its own. Although quite different from those of the Cold War, these problems are still likely to be stress-inducing, perhaps in serious ways. One source of potential difficulty is that the Soviet Union seems on its way toward becoming a modern-day "sick man of Europe," beset by crippling problems and self-doubt. These troubles might translate into an inwardly focused nation, as occurred between 1854 and 1874 when Russia underwent a similar pattern of external defeat and internal turmoil. Conceivably, they could cause the Soviet Union to collapse from within by disintegrating into civil war, with no central government in sight. But equally likely, they could propel a still-existing, still authoritarian Soviet state toward the rigid, suspicious, and bellicose behavior that in 1914 led the Czar to declare war on Austria. Domestic turbulence does not always translate into pacific external behavior, especially among authoritarian nations.

[23]For an appraisal of how the European security system has evolved, see Thomson, 1964.

As of this writing, the Soviet Union's future is clouded. The late 1980s saw major steps, led by Gorbachev, to promote democratization as part of glasnost and perestroika. The Communist Party's role was weakened, power was transferred to state institutions, the USSR's legal system was strengthened, and democratic electoral mechanisms were established. In addition, the USSR began marching toward economic reform, including partial adoption of a free market and other embellishments of Western capitalism. The year 1990, however, saw this reform process deteriorate into decay as the USSR's economy sank into a morass, and numerous republics, including the large Russian republic itself, took steps to withdraw from the Union. In response to this mounting chaos, the political pendulum in the Soviet Union began swinging back toward communist control and authoritarian rule as Gorbachev began distancing himself from the reform movement. In early 1991, Soviet troops were employed to quash secessionist movements in Lithuania and Latvia. Many observers worried that the USSR, its economy spiraling downward and its political system in turmoil, was on the verge of plunging into civil war. For his part, Gorbachev, supported by a cast of increasingly active conservative allies, seemed intent on reestablishing central control even at the expense of the reform movement.

Precisely how this domestic turbulence will play out in the Soviet Union's foreign policy and military strategy remains to be seen. The late 1980s saw the USSR willingly cede its empire in Eastern Europe, embark on the path of withdrawing its military forces, and allow Germany's unification within NATO. In November 1990, Gorbachev signed the CFE Treaty, which called for sharp reductions in Soviet conventional forces, and the Charter of Paris, which called for a new era of democracy, peace, and unity in Europe.

In 1991, the diplomatic pendulum had not yet started swinging backward toward cooler relations with the West, but disturbing signs were beginning to appear. NATO nations openly criticized the Soviet government's use of force against the Baltic republics, coming when the Western coalition launched operation Desert Storm in the Persian Gulf. Worrisome signs began appearing that the CFE and START treaties were in trouble because of questionable Soviet behavior, and conservatives in Moscow began openly criticizing Gorbachev's one-sided concessions to the West. Whether this trend will continue and produce major diplomatic troubles in Europe is uncertain. What can be said is that where a democratic Soviet Union would be prone to cooperate with NATO's nations, a still authoritarian, internally troubled, and disgruntled USSR could be inclined toward a less friendly relationship.

In Eastern Europe, the future is equally uncertain. The dismantling of the Warsaw Pact is likely to unleash a host of pent-up nationalist emotions and long-dormant political conflicts that wracked this region before the Iron Curtain descended. Yugoslavia, Romania, and Czechosovakia all suffer from internal political cleavages that are likely to resurface with the collapse of communist rule. In these nations, as well as in Poland and Hungary, economic troubles magnified by political upheaval as communist rule is abandoned are likely to bring about perpetual instabilities. Moreover, Eastern Europe's troubles will not be entirely domestic. Poland's border problems with both Germany and the Soviet Union, Bulgaria's troubled relations with Turkey, and Hungary's borders with Romania are long-standing sore points. Another serious difficulty looms ahead that nationalist movements within Eastern European nations might spill over to destabilize such USSR border republics as the Baltics, Moldavia, and the Ukraine. This development alone could draw the Soviet Union back into pursuing a threatening agenda in Eastern Europe.

Complicating relations between the Soviet Union and Eastern Europe and continental security affairs generally could be further changes in great power politics. The years ahead are likely to witness the economic and political march of a German-led European Community into Eastern Europe. As the East European nations are drawn into the European Community's orbit, their relations with the Soviet Union might become more distant and strained, especially if democratization and economic reform do not take root in the Soviet Union itself. In that event, the Soviet Union is unlikely to react warmly to the rise in German power brought about by unification, even if Germany is embedded in the European Community and poses no immediate military threat. Soviet-German relations might well become, for good or ill, a central axis of post–Cold War security affairs in Europe.

History suggests that a period of tense German-Soviet relations might lie ahead. Before the 1850s, Czarist Russia pursued an activist diplomacy in Central Europe. Acting with Austria and France, it kept Prussia weak and the German principalities divided. Then in 1856, Russia was defeated by England and France in the Crimean War over control of Turkey. Beset by internal troubles, Russia withdrew into itself and, for the next 20 years, played no active role in continental security affairs. Russia's withdrawal, coupled with Austria's growing weakness, helped pave the way for Germany's unification under Prussia, led by Bismarck.

The Franco-Prussian War of 1870 resulted in Germany's emergence as a dominant Central European power. Propelled by its growing industrial power and nationalist aspirations, Germany began pursuing political influence and economic entanglements in Eastern Europe. This led a suspicious Russia to shake off its internal preoccupation and to reenter European security affairs, a process that began in 1874 when Russia redressed the Crimean War setback by defeating Turkey. From that point forward, a prolonged German-Russian diplomatic struggle ensued, made possible by the weaknesses of Turkey, Austria, and the other East European nations. The resulting tensions between Germany and Russia undermined the political foundations of the balance of power system and helped set the stage for World War I.[24]

This disturbing historical record, in combination with the trends at work today, does not imply that the West will come to remember the Cold War wistfully. At present, nationalism is a less virile and destructive force than it was a century ago. Germany and the Soviet Union currently are pointed toward economic collaboration, not rivalry. Also, the rise of a collective security architecture under the auspices of the Conference on Security and Cooperation in Europe (CSCE) might help tame today's tensions or at least keep them within manageable bounds. But the prospect that these tensions might rise in the years ahead does suggest that the future Europe will be far from conflict-free. Although the Soviet Union no longer will be an ideological enemy of the West, it might well mutate into what might best be called a traditional geopolitical rival. That is, the Soviet Union probably will come to be driven by aspirations and proclivities that, to some important degree, will bring it into cross-purposes with the West, especially Germany. In essence, Europe faces the reappearance, in more muted forms, of a pattern of relations that existed in the late 19th century when a troubled Czarist regime confronted an increasingly confident Germany pursuing its own eastern agenda in an unsettled region. That situation was not a harmonious one.

The prospect of a future European security system with several interacting types of stress, linking local tensions to great power politics, does not mean that war looms ahead.

[24]Thomson, 1964.

Nonetheless, concern about a new form of rivalry between NATO and the Soviet Union, this time separated by a neutral buffer zone, cannot be dismissed from the West's security calculus. War stemming from naked aggression by one side against the other is improbable. But a similarly multipolar system in the late 19th century did produce a long series of diplomatic faceoffs and ultimately World War I. Given this record, there will be a continuing risk in the years ahead that some unforeseen set of interactions might touch off an uncontrollable crisis that escalates, unintentionally, into a military clash. No single scenario may seem particularly threatening, but the combined probability that one of several possible scenarios might occur does seem worrisome.

Given this, the West faces the necessity for choice in determining its policy for managing the European security affairs of the future. In particular, the West will need to address whether the danger of military conflict in Europe is sufficiently high to necessitate continued coalition defense planning under NATO's flag, with sizable force contributions from each member. If so, should the Soviet Union continue to be the focal point of NATO's military planning? In other words, should the West continue to employ a scenario involving war with the USSR as a principal and most demanding contingency for its force planning? Or should it discount the Soviet Union as a potential adversary and design its forces to meet only lesser challenges?

For the first time since the late 1940s these questions have reappeared on the West's policy agenda, and the answers to them are not self-evident. How the West answers them will largely determine the Alliance's future, including NATO's integrated military structure and its conventional posture. Much soul-searching, appraisal, and political debate lie ahead. A broad range of options are open to the West, involving varying degrees of military integration and preparedness. In the end, the West's judgments will be shaped not only by its reading of European trends but also by its attitude toward risk and uncertainty.

Without implying any foreclosing of these issues, this study assumes that the West will adopt a cautious stance toward dismantling its coalition defenses. This does not argue against a phased reduction of NATO's forces as the threat of war recedes and specific requirements diminish, but it does imply that the West will guard against acting precipitously and prematurely. One reason is that Europe is likely to remain unsettled for some time. Another reason is that a basic purpose of defense planning is to buy insurance against the unexpected. This principle might seem unglamorous with peace apparently breaking out in Europe, but it offers a safety net if events do not unfold as many hope. And it is well founded in the harsh lessons of history.

Another assumption of this study is that for some time, the West will continue planning its forces against the contingency of war with the Soviet Union. This does not imply that military conflict with the USSR should be treated with the same expectation that prevailed during the Cold War. However, the Soviet Union will remain a military superpower, and the West cannot afford to reduce its defenses to where they would fail in the unexpected event of a showdown with the USSR. Underlying this assessment, of course, is the assumption that the Soviet Union will remain at least partly a political adversary of the West. Possibly the time will come when democratization in the USSR and the construction of collective security institutions can alter this, but not for the foreseeable future. An inadequately strong NATO defense posture would do more than expose the West risk to a setback in a crisis, it would also weaken the West's ability to maintain a stable balance of power and security architecture in peacetime by presenting the USSR with the kind of military vacuum that creates an incentive for immoderate behavior.

Finally, the assumption also is made here that although the Alliance's military strategy will undergo important changes, NATO will need to continue relying on a strong conventional defense capability. Extended nuclear deterrence will need to remain a cornerstone of NATO's doctrine, but any shift back toward a largely nuclear strategy would be both unwise and unnecessary. It would also be politically unacceptable in light of the growing public aversion to nuclear weapons in Europe. Moroever, the impending military reductions in the Soviet/Warsaw Pact military posture will make conventional defense a more feasible NATO goal than at any time in the past two decades. Indeed, the objective of a stalwart (confident and enduring) conventional defense is now within reach.

The Military Regime Ahead

Assuming that NATO will continue focusing its defense planning on the Soviet Union, what are likely to be the military dimensions of the European security environment ahead? In a nutshell, the future Soviet "threat" (military capability) in Central Europe will be less daunting than during the Cold War by an appreciable margin, but it still will be a matter to be taken seriously. Three developments will pare the traditional threat back in both size and timing: the departure of the East European nations from the Warsaw Pact, the withdrawal of Soviet forces from the forward areas, and the asymmetric cuts imposed on the USSR by the CFE accord. These already underway changes are likely to continue unfolding in stages, over several years. As a result, the West will be facing a complex transition process in Central Europe, each phase of which is likely to pose different implications for NATO's force posture and employment doctrine.

These changes are best understood if they are viewed separately. Table 3 displays how the defection of the East European nations will reduce the Soviet/Warsaw Pact force posture that NATO traditionally has planned against in Central Europe. The table shows remaining force levels for the three major contingencies that have occupied NATO planners. Forces for a short warning attack at M+4 are reduced by one-third. Even more noteworthy, forces for an attack with forward-deployed Warsaw Pact forces at M+8 are reduced by over 50 percent. In comparison with NATO's current posture, force ratios in Central Europe would be close to parity. The reduction for a fully mobilized Soviet attack at M+30–45 is large, but less militarily decisive. By reinforcing from the western military districts and other districts in the

Table 3

Estimated Soviet/Warsaw Pact Force Levels in Central Europe (DEs/DEFs)

	M+4	M+8	M+30/45
Original posture			
Soviet	27	27	59–87
NSWP	13	31	31
Total	40	58	90–120
Post-withdrawal of East European nations			
Soviet/total	27	27	59–87

rear, the Soviets would still be able to assemble a purely Soviet force that is large enough to meet traditional planning standards for a Central Region offensive.

This reduction will be further magnified by the November 1990 signing of the CFE accord, whose effects in Central Europe are displayed in Table 4. Implemented over the coming five years, the CFE accord will impose at least a 50 percent cut in Soviet force levels in both the forward areas and the USSR's military districts, encompassing the entire region from the Atlantic to the Urals (ATTU). Soviet forces in Eastern Europe will be reduced to the equivalent of about 12 combat divisions. The forces in the USSR's military districts that traditionally were thought of as contributing to the Central Region will be reduced to about 28 divisions. Other military districts in the ATTU region will house another 17–24 divisions. Of the total in the ATTU region in the post-CFE era, about 40 divisions could be made available for a Central Region war.

The process of reduction will be completed by the withdrawal from the forward areas of Soviet forces, already scheduled to be withdrawn from Hungary and Czechoslovakia by late 1991. Withdrawal from eastern Germany will occur by 1995, leaving the Soviets entirely without a forward presence that could have permitted a short-warning attack. Counting all available forces in the ATTU region and allowing for defensive requirements elsewhere, the Soviets would be able to assemble only about 30–40 divisions for offensive use in the Central Region. Plausibly other forces could be drawn from outside the ATTU region (Siberia, Southwest USSR, and the Far East) but only within certain limits and over a more extended period of time. Table 4 displays the combined effects of CFE and complete Soviet withdrawal.

Table 4 thus demonstrates how favorably (for NATO) the changes underway are likely to alter the military balance in Central Europe. For all practical purposes, the defection of the East European nations and implementation of the CFE accord will end the threat of a short-warning attack as long as NATO's residual posture is maintained near present levels. The USSR will retain a capacity to employ Soviet forces alone, but in less quantity and after a longer period of mobilization and reinforcement.

This drawdown in Central Europe will be taking place against the background of an overall change that will sharply truncate the USSR's ability to launch a global offensive campaign against the West's interests. Table 5 displays the combined effect of these three changes on the total Soviet/Warsaw Pact posture. During the Cold War, Western estimates held that the Warsaw Pact's 267 total divisions provided the Soviet Union the capability not only to attack Europe's Central Region and the flanks but also to conduct an offensive concurrently in Southwest Asia and to exert pressure on U.S. allies in Northeast Asia. The Warsaw Pact's collapse and CFE cuts will leave the Soviet government with less than one-half

Table 4

Soviet Forces vs. Central Region
(DEs/DEFs)

	M+4	M+8	M+30/45	M+60/90
Original posture	40	60	90–120	90–120
Post NSWP defection	27	27	59–87	59–87
Post CFE (Soviet only)	12	12	40	50–65
Post Soviet withdrawal	0	0	25	50–65

Table 5

Combined Effect of Reductions on Soviet/Warsaw Pact Posture (DEs/DEFs)

	Traditional Posture	Post-CFE Posture	Available for Central Europe
Soviet			
Forward areas/ western MDs	68	30–35	30–35
Other ATTU MDs	73	30–35	15
MDs outside ATTU	66	40–50	5–15
NSWP[a]	60	0	0
Total	267	100–120	50–65

[a]Non-Soviet Warsaw Pact.

this total capability: 100–120 divisions. As a result, the Soviet Union will no longer be capable, minus a lengthy (e.g., two year) mobilization that creates additional forces-in-being, of a multipronged offensive on several different axes of advance. This change will considerably ameliorate the threat of a global, or at least Europewide, Soviet offensive that bedevilled Western planning for many years. This judgment, of course, is sensitive to future Soviet force levels. A smaller Soviet posture than portrayed here would truncate the USSR's options even further. A larger posture—as could come about if the USSR retains much of the equipment withdrawn from the ATTU under the CFE treaty—could broaden the USSR's options.

Notwithstanding this expected drawdown, the Soviet Union will still remain a military superpower with major offensive options in Europe. The USSR will continue to possess strategic nuclear weapons that could be used against Western Europe and that will need to be counterbalanced by continued U.S. extended nuclear deterrent coverage over NATO's members. Also, the Soviet Union, while no longer able to quickly mobilize and to attack concurrently on several geographic fronts, will still be able to engage in offensive operations on a more limited, regional basis. Within a fairly short time (weeks and months, not years) it will have the ability to assemble perhaps 50–65 divisions to pursue a limited but politically threatening offensive in Central Europe. Simultaneously, it might also be able to exert at least limited pressure on at least one nation along NATO's flanks (e.g., Norway or Turkey). Regardless of the USSR's intentions, about which NATO is uncertain, the existence of a physical Soviet military capability to threaten Europe will not become a thing of the past.

The exact dimensions of this physical capability will depend heavily on how the Soviet Union structures its forces in the years ahead. The USSR will have a broad range of choices at its disposal, choices that could pull its armed forces in either a defensive or an offensive direction. A CFE accord will allow the Soviets to deploy some 12,000 tanks in the ATTU region, along with a sizable number of artillery pieces, armored troop carriers, helicopters, and tactical aircraft. All told, the Soviets should have enough equipment for about 50–60 heavy tank and motorized rifle divisions at likely unit inventory levels. Conceivably, additional weapons could be stored in prepositioned stocks east of the Ural mountains and earmarked for units in the ATTU region, thereby adding to the total.

Evidently, additional Soviet divisions in the ATTU region will be structured as artillery-rifle (infantry) units. Ostensibly intended for positional defensive warfare, these units will not have the tanks that are normally associated with offensive operations. But infantry units are not entirely lacking in cross-country mobility if they are motorized, and they can be incorporated usefully into a larger posture that contains armor elsewhere. The German Wehrmacht in World War II, after all, was largely composed of infantry units. In the design of an offensive posture, the critical issue (beyond certain limits) is not tank and artillery inventories, but rather how the combined arms team is assembled. For years, the Soviet Army has been so heavily equipped with tanks and artillery that it did not have a balanced combined arms posture. If the Soviet Army reacts to CFE's limitations by strengthening itself in other less-endowed areas (e.g., infantry, logistics, and C3I), it might preserve a better offensive capability than seems likely at first glance.

Thus far, Soviet forces seem to have been suballocated on a regional basis across the USSR. Forces deployed in Central Asia and the Far East, consequently, have not been regarded as being oriented toward Europe. As Soviet units in the ATTU region are reduced, it might become necessary to reorient non-ATTU forces to Europe, where they previously were not needed. Equipment stored east of the Urals could be used in a crisis to augment the Soviet Army in a variety of ways. Although the distances involved are long, the USSR has a well-developed rail network. Relations with China and other neighboring countries are a potential constraint, but they are not absolutely paralyzing, since the USSR possesses a nuclear deterrent. Physically, the USSR could copy the U.S. practice of treating its entire CONUS[25]-based posture as a mobile reserve that could be deployed anywhere.

Future Soviet military capabilities in Europe and elsewhere will be influenced heavily by the goals set for modernization and readiness. In past years, the large size of the Soviet Army and Air Force resulted in an uneven investment pattern in both areas. The most important army units (e.g., the GSFG) were equipped with modern weapons and were fully staffed in peacetime. Less important units, especially the numerous forces deployed in the USSR's rear districts, were given older equipment and were staffed with only a peacetime active cadre. These units were less ready and had less potential combat power than their more well-endowed counterparts. In the years ahead, financial constraints and internal political troubles could conceivably result in even lower overall modernization, readiness, and sustainability. In that event, the Soviet Army might deteriorate into a hollow force that cannot quickly generate major combat power. A quite different trend, however, also is possible. The reduced size of the Soviet Army might facilitate a concentration of funds on a smaller number of remaining units, resulting in a smaller but more uniformly equipped and well-trained army that is effectively prepared to execute Soviet war plans. The United States followed exactly that course in the early 1960s, when the number of U.S. Army divisions declined but overall combat power actually increased.

The West's uncertainty about future Soviet military capabilities for waging conventional war is unlikely to be resolved anytime soon. The key issue is whether the USSR, in the years ahead, will aspire to retain at least a regional counteroffensive capability with respect to Europe. Historically, the Soviet Union and Russia before it have always had large armies capable of mobilizing quickly and entering Eastern Europe to defend Russian national interests. Napoleon and others discovered this capability to their chagrin. The evidence of history and the USSR's own long and exposed borders argue against the likelihood

[25]CONUS = Continental United States.

that this step will be taken anytime soon. Future Soviet decisions in this area clearly will be a factor weighing heavily in the West's own defense calculus.

Even if the USSR does elect to preserve a regional offensive posture, the Soviet threat to Central Europe will unavoidably be reduced in size and shifted outward in time. Figure 8 displays how these changes will transform Soviet time-phased mobilization, reinforcement, and buildup capabilities in Central Europe. The defection of the non-Soviet Warsaw Pact (NSWP) nations will prevent the Soviets from mounting a major invasion of 40–60 divisions within the first week; the best that could be done is a 25-division operation with only Soviet forces. By drawing heavily on USSR-based reserves, the Soviets still could assemble an 80-division posture, but the process would take several weeks longer. NATO would face a large invasion but would benefit from additional time to make preparations of its own.

Figure 8 further shows that a CFE agreement will effectively deny the Soviets a full invasion posture even with additional buildup time. Under CFE, the Soviet presence in the forward areas is reduced to only about 12 divisions. Equally important, the Soviets would be able to assemble a total force of only about 50–65 divisions after drawing on the western military districts and other districts within and outside the ATTU region. In absolute terms, this invasion force still would be a potent one; but comparatively, it is smaller and more manageable than the traditional Warsaw Pact threat. Operationally, the Soviets would no longer be capable of building the sizable second strategic echelon that has been a particular source of worry for NATO. In purely military terms, it is a CFE agreement, and not the Warsaw Pact's collapse, that physically hamstrings Soviet invasion options. The present Soviet Army is so large that the East European forces are not needed to threaten NATO. For this reason,

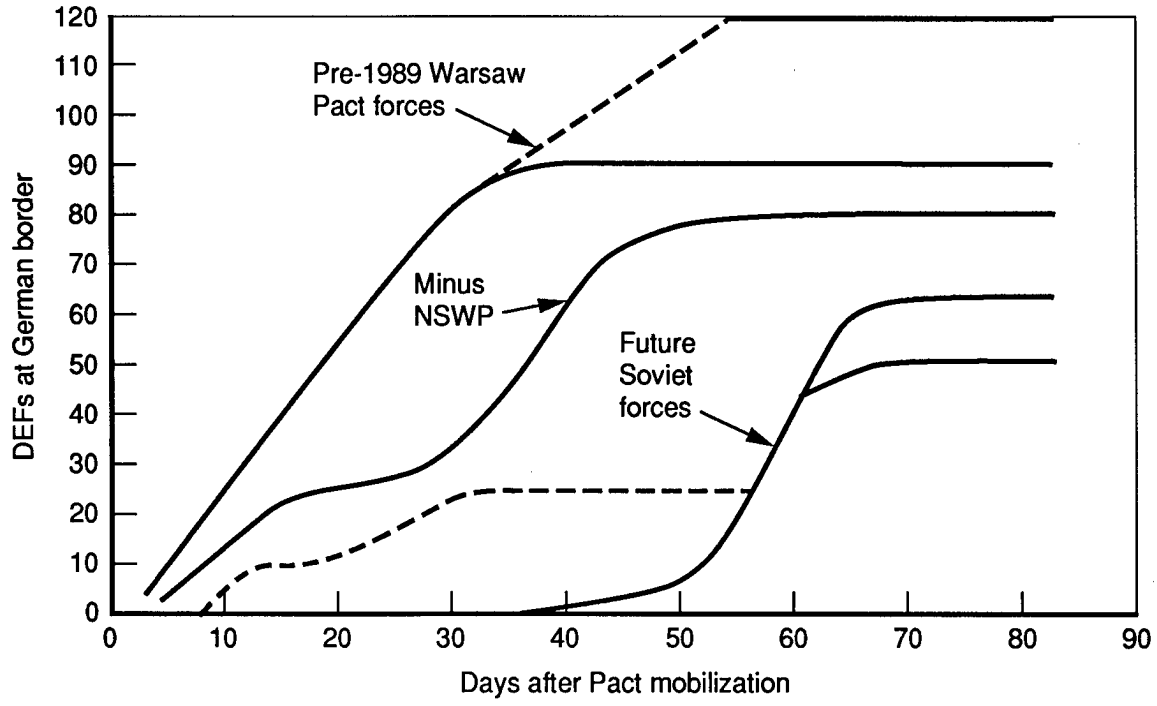

Fig. 8—Future Soviet Buildup Rates

failure to implement the CFE treaty could leave NATO facing a full Soviet invasion capability even if Soviet forces are withdrawn from the forward areas.

Toward Parity, or Something Like It

These reductions will have a major influence in favorably transforming the NATO/Soviet force balance in Europe, provided NATO itself retains a defensively capable posture. Figure 9 illustrates the magnitude of these potential changes by displaying time-phased comparative force ratios near the German border, assuming both sides begin mobilizing on the same day. The figure is not based on a crisis scenario that accounts for possibly different M-days, it merely measures physical potential on both sides. Its central message is that after CFE is implemented, the Soviet Union no longer will be capable of achieving the 2:1 edge in ground forces that has given it a potentially dominant military position in Central Europe. Once Soviet forces are withdrawn, by definition NATO will possess greater peacetime forces there. After mobilization and reentry, the Soviets could assemble enough forces to match NATO's forces and possibly exceed them by a margin of 1.4:1. But this margin would be far less than the previously overpowering difference of 2:1. NATO would face a situation of close to parity.

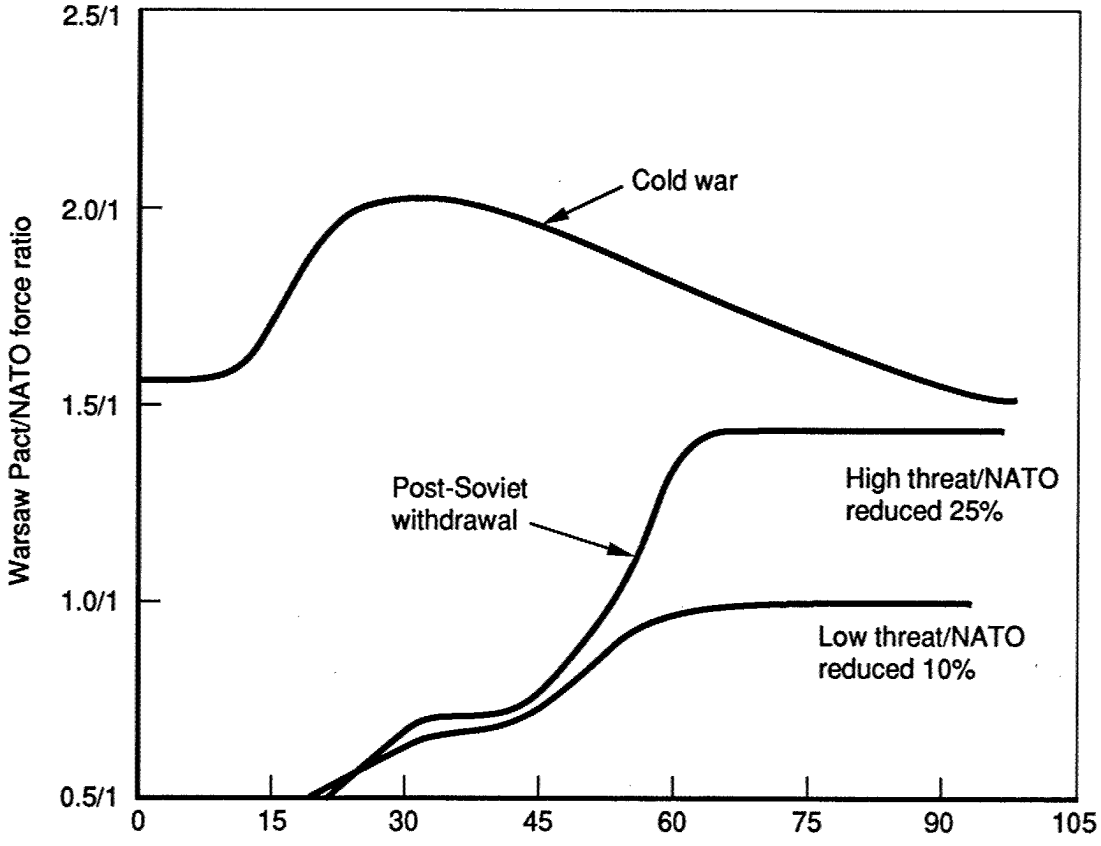

Fig. 9—Time-Phased Force Ratios in Central Europe

Near parity, if actually attained, will stabilize the Central Region force balance and enhance NATO's defense prospects. However, this rosy picture could become clouded by actions on NATO's part in adjusting its own force posture to the less threatening environment of the post–Cold War era. Reductions of 10–25 percent are being openly discussed in several NATO capitals, including Washingon, D.C. Some observers are calling for 50 percent cuts, albeit normally in the context of a CFE II accord that further reduces Soviet forces by a similar amount. Large unilateral NATO cuts could undermine some of the benefits of the Soviet/Warsaw Pact drawdown.

Even reductions in the 10–25 percent range will compel NATO to rethink its linear defense philosophy. Parity is no guarantee that NATO can defend successfully. History shows numerous cases of ineffective defending armies being beaten at parity. A sound theater employment doctrine is sine qua non for a successful defense effort against a skilled attacker. The problem facing NATO is that sizable force cuts will render its linear defense vulnerable, even at near parity, by stripping away its already scarce reserves. This prospect does not doom NATO's conventional defense strategy, but it will compel NATO to craft a new employment doctrine that can be executed with available forces.

Concern that the Cold War's end is also bringing about the demise of linear defense is one powerful reason why NATO military authorities and some outside observers have begun to discuss a different theater employment doctrine. Considerable uncertainty exists about exactly at what point of reductions the layer cake becomes vulnerable. The subject is complex, and it is difficult to identify a single threshold below which linear defense is impossible. But if the layer cake were maintained in some rigid and unaltered form, the military risks to NATO evidently would increase as Alliance force levels decline. An emerging Western consensus on this point alone would probably have propelled NATO in the direction of an altered employment doctrine even without the political revolution in Europe of 1989.

Unification of Germany

Adding a powerful political impetus to this trend is the disappearance of the German Democratic Republic and the emergence of a unified German republic. Unification was not on Europe's political agenda during early 1989 when the military debate over NATO's employment doctrine began building. The dominant assumption in technical circles therefore was that the inter-German border would remain the geographic anchor of NATO's defense strategy. This assumption was shattered when the Berlin Wall came down in November that year, and a few weeks later Soviet leader Mikhail Gorbachev stunningly endorsed the idea of unification. Gorbachev attached important qualifications to his stance, but his action set in motion forces of change that promised to reshape the fundamentals of the European security system, including the locations of the borders that have defined NATO's forward defense strategy.

During early 1989, Soviet spokesmen continued to insist that Germany's detachment from NATO was the price the West had to pay for unification and the departure of Soviet troops from Eastern Germany. To replace NATO and the Warsaw Pact, the Soviets offered to help create a European Security Council that, along with CSCE, would establish a collective security regime in Europe. The West, including the United States and the FRG government itself, firmly rejected this position. Citing the stabilizing effects of keeping Germany in the Atlantic Community, it insisted that a unified Germany remain as a member not only of NATO, but also of its integrated military command. For a time, the two sides seemed headed for a showdown in the talks being convened to set the terms for unification.

Unfolding events in Europe soon overtook this debate. Germany's headlong rush toward unification, including the victory of the Christian Democratic Union (CDU) in East Germany's elections in March, made it evident that Socialist rule there was dead. Further weakening the USSR's bargaining position came the collapse of communism across Eastern Europe, accompanied by Moscow's acceptance of agreements to withdraw its forces from Hungary and Czechoslovakia by 1991. Emboldened by these developments, the West held its ground on Germany's staying in NATO, but it did offer a set of concessions that were designed to assuage Soviet worries. Included were the Genscher Plan, which promised not to deploy NATO combat forces into eastern Germany, and FRG reassurances on the inviolability of the German-Polish border. Also, NATO's foreign and defense ministers expressed willingness to initiate discussions on reduction of short-range nuclear weapons in Europe. Further, the West began warming to the idea of expanding CSCE's role, provided that NATO remain a part of any collective security regime.

The Soviet government publicly reacted to these gestures by remaining adamant through the initial session of the 2+4 talks and the Washington summit in late May. But in the background, unofficial signs of growing Soviet flexibility began to appear, including Gorbachev's admission that Germany would have to be allowed to choose its own destiny. By the time 2+4's second session was convened in late June, the Soviets appeared to be bowing to the inevitable. The London Summit evidently sent Gorbachev the conciliatory gestures that he needed to take the final step. Shortly thereafter, German Chancellor Helmut Kohl and Gorbachev met in the Soviet Union, and there Gorbachev formally agreed that a unified Germany could remain within NATO. The talks were brought to a successful conclusion in September when a peace treaty was signed that ended allied occupation rights in Berlin. Germany formally unified on October 3, 1990.[26]

Unification means that the territory of eastern Germany, as part of the Federal Republic, will automatically qualify for defense protection under the North Atlantic Treaty. Articles 5 and 6 of the Treaty commit the Alliance members to protect "the territory" of its members. Also, West Germany's entrance into NATO in 1955 was conditioned on the promise of Alliance support for unification, when that time came. If the German government seeks NATO defense coverage for its new territory, the Alliance will be obligated to extend protection to eastern Germany. That will be the case whether NATO's other members approach the task eagerly or reluctantly. It is a natural extension of the transatlantic bargain that forms NATO's foundation.[27]

Unification is unlikely to bring about any wholesale movement of NATO forces eastward. The Treaty on the Final Settlement with Respect to Germany, signed in September 1990, agrees not to deploy forces assigned to NATO (German or otherwise) east of the old inter-German border. Nonetheless, unification will compel NATO's defense policy to begin focusing eastward. Although peacetime deployment of NATO's forces into eastern Germany is not practicable, wartime employment there is another matter entirely, especially if that area were subjected to Soviet aggression in a crisis.

NATO's wartime employment plans to include defending eastern Germany will broaden gradually. Most probably, they will evolve in parallel with unification dynamics and with the phased withdrawal of Soviet forces. The entire process, nonetheless, seems destined to challenge the very fundamentals of NATO's military doctrine and the diplomatic arrangements that support it. Over the course of the Cold War, the concept of defending only along the

[26]*Washington Post*, "Gorbachev Drops Objection to United Germany in NATO," July 17, 1990; "German Merger Clears Hurdles," July 18, 1990.

[27]See NATO Information Service, *The North Atlantic Treaty Organization: Facts and Figures*, Brussels, 1984.

inter-German border has sunk deep roots into NATO's military programs and internal political understandings. To one degree or another, these roots will now have to be pulled up and replanted.

The Larger Setting

The task ahead for NATO will be one of designing a new employment doctrine that not only replaces linear defense but also defends eastern Germany. A host of complicated considerations, strategic and tactical, are involved. Suffice it here to say that the Alliance seems to be entering into a period of reappraisal unlike anything experienced since the 1960s, when the basics of NATO's military strategy were being crafted.

The problem of defending a unified Germany will need to be addressed within a much larger setting, one that is likely to impose new and different constraints on NATO's military strategy. Although NATO's defense policy has concentrated on Central Europe, the northern and southern flanks remain important regions, and will present challenges of their own in the coming decade. In particular, instability is likely to manifest itself most openly in the Balkans, where Soviet and Western interests intersect in ways that historically have given rise to tensions and periodic crises. NATO evidently will continue to require a deterrent posture built along classical lines for defending Turkey, Greece, the Mediterranean, and Western interests in the Middle East. But NATO is also likely to require the capacity to act in new and more flexible ways, guided by the dictates of post–Cold War diplomacy. One possibility, for example, is the deployment of Western forces into Yugoslavia and Romania in peacekeeping roles. Requirements of this sort would draw Western nations into the larger task of fostering stability across all of Europe and into areas beyond the normal purview of NATO's military strategy.

Within Central Europe itself, defense issues will probably command less saliency, urgency, and credibility. The fading of the Cold War and the withdrawal of Soviet forces, however, do not render these issues unimportant. Above all, the new role of military forces in shaping the Europe security order will need to be understood and incorporated into NATO's defense policy, including its theater employment doctrine.

Unless the trends now at work are somehow reversed, the future European military balance seems destined to be less volatile than during the Cold War. NATO no longer will be confronted with the threat of being outnumbered and overpowered by a tightly integrated adversary alliance. Soviet and NATO forces will be separated by lengthy distances and no longer poised to strike at each other's throats on short notice. The level of armaments is likely to be smaller, and forces on both sides will be maintained at lower levels of readiness. These developments will enable NATO's military strategy to aim for a more stalwart conventional defense and to reduce its traditional high reliance on nuclear weapons. More fundamentally, they cannot but help reduce tensions in Europe, ameliorate competition in armaments, foster greater crisis stability, and promote better overall political stability.

At the same time, defense issues are likely to remain alive, thereby constraining NATO from fully shedding its military trappings as it pursues the path of a "more political" alliance. In this regard, it should be remembered that the physical reduction of armaments itself does not guarantee peace. Politically decisive wars still can be fought with much smaller forces than the twentieth century has witnessed in Europe. The Franco-German War of 1870 was fought with armies of about 400,000 on each side, roughly one-fifth of the forces that NATO and the Warsaw Pact have been planning to field. Waterloo, Austerlitz, and Gettysburg were

fought with less than 100,000 soldiers on each side, about the equivalent of a single modern corps. Especially since modern technology has dramatically increased the firepower of today's armies, small forces can still be used as instruments of national diplomacy.

In the years ahead, NATO's employment doctrine will be a factor in the stability equation. NATO will need to design a doctrine that responds to the emerging military balance of near parity at lower force levels. But NATO's future doctrine, linear or otherwise, must reflect more than military considerations. It also must respond to NATO's political goals as they come to be interpreted in the post–Cold War era, including the objectives of Alliance unity, control of escalation, and forward defense. The challenge facing NATO will be to fashion a theater employment doctrine that makes both military and political sense.

3. TOWARD A NEW NATO THEATER EMPLOYMENT DOCTRINE

This section addresses the issue of how NATO should design its future theater employment doctrine in Central Europe to meet the Alliance's shifting security requirements in the years ahead. Its analytical approach is prescriptive, not evaluative, of an existing plan. Rather than suggest some fixed blueprint for NATO to follow it constructs an analytical framework that helps identify the alternatives open to the Alliance and evaluates these alternatives in a cost-benefit fashion.

For military reasons alone, NATO will need to determine whether a better employment doctrine can be found to replace its traditional layer cake array and linear defense.[1] This will be the case whether or not NATO's defense plans are shifted eastward from the old inter-German border to provide protection for all of a unified Germany, and if so, when. The task of designing a coherent doctrine will therefore remain an important item on NATO's agenda for as long as balancing Soviet military power in Central Europe continues to be a key feature of Alliance military strategy.

The need to replace a timeworn and vulnerable doctrine does not guarantee that a better one readily can be found to take its place. Designing the old layer cake was a complex enterprise, one that required the balancing of complicated political and military tradeoffs. It was not accomplished overnight. Multiple considerations, often pulling in different directions, are also likely to enter into the design of a new NATO employment doctrine.

What kind of theater employment doctrine makes better sense than simply maintaining the old linear defense and moving it eastward at the appropriate juncture? What are the important criteria, costs, benefits, tradeoffs, and other factors involved in designing a new doctrine? And what are the implications of a new doctrine for NATO's force needs, program priorities, and internal distribution of national missions? Can NATO hope to continue doing its conventional defense business as usual at lower force levels and preparedness standards than before because the Soviet military threat has been reduced? Or will the act of designing a new employment doctrine impose a need to undertake major innovations in its traditional practices?

Although designing a better theater employment doctrine is not easy, neither is it a fruitless enterprise. Provided the Alliance is willing to face the demanding challenges created by innovating in this area, it is possible to reduce the risks that a linear defense at low force levels would bring. A three-fold sequence of doctrinal changes could evolve in tandem with the withdrawal of Soviet forces from eastern Germany.

- In the near term, a large Soviet military presence will remain in eastern Germany. During this period, NATO's employment doctrine would still defend in the vicinity of the old inter-German border, but it would do so in more flexible and adaptive ways than in the past to strengthen its capacity to counter-concentrate faster and more effectively. NATO would shift forces northward and adopt such practices as thinning out its forward commitments to create more reserves, trading space for time, and conducting strong counterattacks. NATO's doctrine thus would become less linear, less positional, more mobile, and more maneuver-oriented.

[1]The complex military reasons why a linear defense is difficult at low force levels are addressed, in technical and historical detail, in Sec. 4.

- In the mid-term (a transitional period through 1994–95), only small Soviet forces will still be stationed in eastern Germany. With less need to fear a surprise Soviet attack in a crisis, NATO's wartime employment doctrine would begin to move eastward. In particular, it would concentrate on remedying the traditional vulnerability of NATO's concave defense line by establishing defense positions on such favorable terrain features as the Elbe and the Salle rivers.
- The final stage (end game) will be reached once Soviet forces are withdrawn by 1995. At that juncture, NATO's defense umbrella would be extended eastward to the Oder-Neisse rivers on the German-Polish border. NATO's employment doctrine then would be neither linear nor positional. Rather, it would be one of an "echeloned" defense. Strong forces would be committed on the Oder-Neisse line, where an enemy attack probably would be launched, but they would defend in flexible ways. Small forces would be planned for defense of the less-threatened German-Czechoslovakian border region. Meanwhile, large NATO armored and mechanized forces would be deployed in the rear areas, from where they could launch counterattacks against enemy salients.

Movement toward doctrinal concepts of the sort envisioned here would require major changes in NATO's traditional ways of doing business. NATO's forces would have to learn how to mobilize differently and to fight differently than in the past. The entire process would pose new requirements for NATO's weapon systems, combat forces, support structures, and standards for readiness and sustainability. Additionally, there would be a need to realign national responsibilities away from the philosophy that committed several different nations to the forward areas. A new mix of national missions and a greater emphasis on multinational formations for combat and support forces could be disruptive and difficult to implement, but the Alliance would be able to achieve a stalwart conventional defense posture for the first time in its 40-year existence and truly render nuclear weapons an option of last resort.

FACTORS AFFECTING NATO'S DECISION CALCULUS

In principle, NATO does have the option of retaining the old layer cake linear defense. It could simply maintain the current doctrine while Soviet forces are still based in Germany and then move it eastward after they have departed. That would follow the path of least resistance and avoid a stressful reexamination of military policy. However, NATO's employment doctrine then would be of declining relevance to the conditions ahead.

NATO's Future Force Levels

One barrier to perpetuating the old doctrine is that NATO's forces evidently will fall well below present and even post-CFE I levels. The Alliance already has committed itself to a follow-on CFE II negotiation. Although NATO's initial statements on CFE-II's mandate do not address the issue of further reductions in equipment levels, they do envision constraints on offensive capability, manpower levels, readiness standards, and the size of the German Army. Out of these constraints could easily come further restrictions on the overall size of NATO's conventional defense posture.

Quite apart from CFE's future, most NATO governments already are planning unilateral force cuts below what originally had been envisioned as CFE I's outcome. The United States plans to reduce the size of its active army by up to six of the present 18 divisions. The German government plans to scale back its active army from 12 to nine divisions. The British have announced a 25 percent cut on total military personnel, a 50 percent reduction of the British Army on the Rhine's (BAOR) manpower, and withdrawal of at least one division from Germany. The Belgian, Dutch, and Danish governments are already planning similar reductions. The French defense posture, beset with fiscal constraints arising from expensive modernization programs, seems unlikely to escape reductions of some magnitude.

These trends do not mean that NATO's members are on the verge of disarming themselves. The London Summit produced official recognition that the Alliance will still have to retain sizable forces in Central Europe. Nonetheless, military budgets clearly are drifting downward. Most national defense ministries appear to be fastening onto programmatic strategies aimed at keeping their modernization accounts fairly intact while cutting force levels and readiness standards. These cuts are likely to fall most heavily on active army formations, which are expensive because of their high personnel and operating costs. Many of the deactivated units evidently will be transferred into the reserve structure rather than be disbanded altogether. In theory, these units would be available after the more prolonged period of warning that is likely to accompany future European scenarios. But even so, NATO's future conventional defense posture seems destined to be smaller and less powerful than was the case during the 1980s.

Exactly how far NATO's posture will shrink is uncertain. A reasonable estimate is that reductions of 10–25 percent are forthcoming by the time Soviet forces have left Germany. Table 6 displays this estimate in terms of NATO's buildup rate for ground forces (DEFs) and tactical combat aircraft during the first 75 days of the mobilization and reinforcement process. This table is an estimate of future force levels, not a statement of requirements for executing any particular employment doctrine. It assumes that U.S. peacetime deployments are reduced to 2-1/3 Army divisions and four tactical fighter wings. The U.S. contribution at M+30 is assumed to be ten DEFs and 1200 combat aircraft. Total U.S. deployments by M+75 mobilization are assumed to be 17 DEFs and 1200 combat aircraft, 25 percent below today's total. The German Army is reduced to nine active divisions and, counting infantry reserves, 15 mobilizable DEs. When CFE I equipment levels are taken into account, a unified Germany will be capable of deploying about 13 DEFs after mobilization. Other Allied forces are reduced equivalently. For all periods, French forces are assumed to be committed only by M+15–30 and therefore are not counted in the category of "peacetime active" forces. Total active Allied forces are reduced by about one-third; totally mobilizable Allied forces, counting reserves, are reduced to some 28 DEFs, or by about 15 percent compared with the situation in 1989.

Table 6 suggests that NATO's forces in Central Europe by 1995 are likely to be reduced well below the levels normally associated with a linear defense, particularly in the buildup process. The situation would begin to change with mobilization and outside reinforcement, but by 1995, even NATO's fully mobilized posture would be hard-pressed to execute a linear defense against an enemy threat as large as 65 divisions.

Table 6

An Estimate of NATO's Future Buildup Rates in Central Europe (DEFs)

	Peacetime Active	M+15–30	M+60	M+75
1989				
DEFs	24	45	54	57
Combat aircraft	1,700	3,600	3,600	3,600
1992–1993				
DEFs	21	41	48	51
Combat aircraft	1,500	3,200	3,200	3,200
1995 and beyond				
DEFs	17	37	42	45
Combat aircraft	1,200	2,800	2,800	2,800

Changing Strategic Conditions

The need to develop a new doctrine is reinforced by the changing strategic conditions that lie ahead. NATO will confront three stages of security affairs over the next five years, each of which calls for a new employment concept. For the next two years or so, the Soviets will retain sizable forces in eastern Germany, along with the infrastructure for quickly absorbing large outside reinforcements. This presence will enable them to assert de facto military control of the eastern German terrain in a crisis. As a result, they in theory will still be capable of launching an attack from their present bases into western Germany, but only temporarily so.

The years 1993–94 will mark an important transition stage away from this pattern. Soviet force levels and infrastructure in Germany will have declined to the point where the Soviets no longer will be able to assert physical control over eastern Germany in a crisis. They also will be heavily outnumbered by NATO's forces based in western Germany. At the onset of a crisis, Soviet forces remaining in eastern Germany would probably have to retreat initially into a defensive enclave around Berlin and behind the Elbe River and remain there until reinforcements arrived from the USSR. Only then could they undertake offensive operations. A Soviet attack would have to be launched from this enclave and then cross eastern Germany before the old inter-German border could be reached.

After 1994–95, the end state will be reached. Soviet forces will have completely withdrawn from eastern Germany and retreated behind the USSR's borders. Any Soviet offensive operation would have to spring forward from the USSR and travel across Poland before Germany could be threatened. All of eastern Germany will be free of any peacetime Soviet military control and therefore could have NATO's protective security umbrella extended over it. From that point forward, NATO will be in a position to contemplate defense of eastern Germany all the way up to the Oder-Neisse border with Poland.

NATO's conventional defense strategy will need to evolve in a parallel three-stage pattern. In the near term, NATO plans will need to remain concentrated on defense operations in the vicinity of the old inter-German border. During the transitional mid-term, NATO's strategy will be able to contemplate prospects for moving at least part way into eastern Germany. From the end state onward, NATO's plans will need to address the task of defending

all of eastern Germany from a reconstituted Soviet threat crossing Poland. In essence, NATO's defense focus will be able to extend eastward as far as the Oder-Neisse rivers. By definition, these changing strategic conditions will compel NATO to alter the geographic location of its defense line by moving it progressively eastward. The Alliance can hardly afford to continue defending along the old inter-German border once NATO's forces are physically capable of moving into eastern Germany in a crisis.

The impracticality of preserving NATO's layer cake becomes evident when the implications of applying it in eastern Germany are considered. Figure 10 displays how a NATO linear defense along the Oder-Neisse line and the German-Czechoslovakian border would be formed if the old layer cake were to be transposed eastward. It assumes a literal transposition, with no changes in the nature or geographic mix of national corps sector responsibilities.

Even at first glance, imposing political and military problems become evident. A layer cake along the Oder-Neisse line would require that both Belgium and the Netherlands continue accepting forward corps sector missions. Whether these nations would be willing to do so is rendered uncertain by the extent to which the battle zone will have been moved away

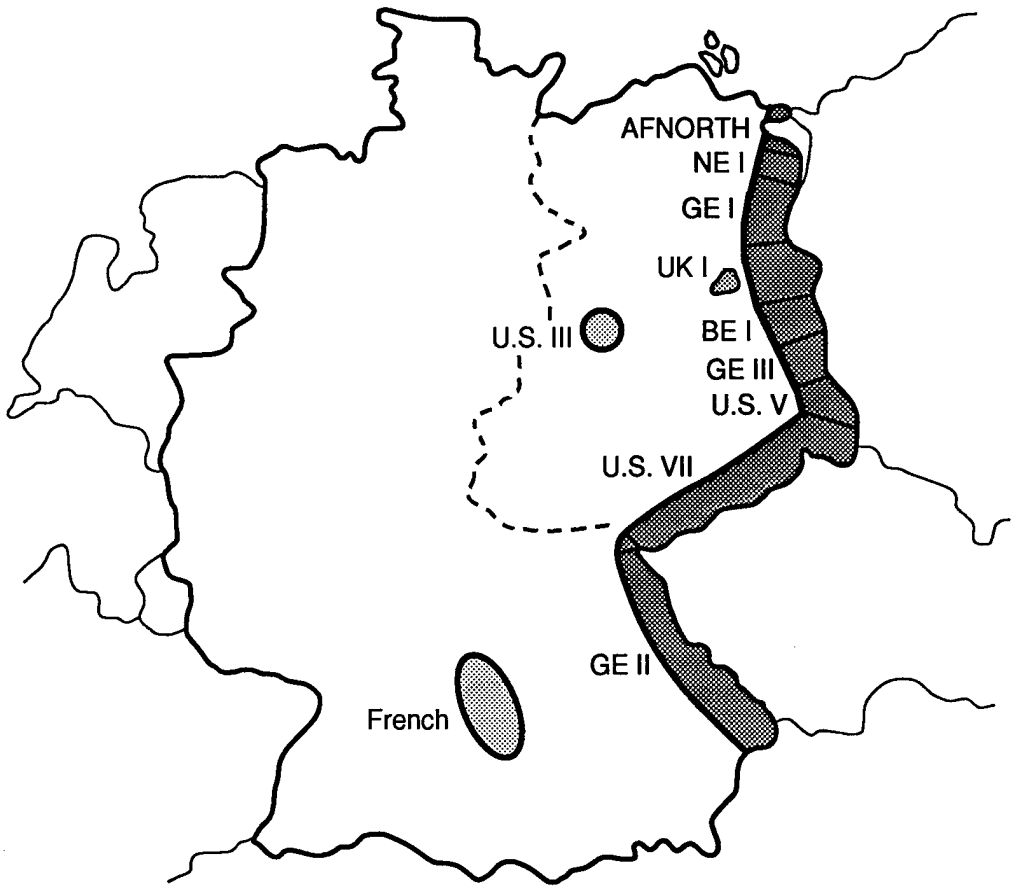

Fig. 10—NATO's Layer Cake Along the Oder-Neisse Line

from their own borders. Quite apart from their political intentions, serious questions would arise regarding their logistic capability to support combat forces whose wartime deployments had extended their supply lines by an additional 400 kilometers.

Other problems would arise. British and Belgian forces would acquire the mission of defending Berlin, Germany's capital. Only one German corps would be deployed along the entire Oder-Neisse line, the focal point of a Soviet attack. Meanwhile, the remainder of the German Army, along with most U.S. forces, would be deployed well to the south. Their mission would be to defend the lengthy German-Czech border, amounting to 480 kilometers of frontage. They doubtless could accomplish this mission but in ways that would extend beyond the needs of the situation, since Soviet forces probably would not attack there heavily. The net effect would be to divert key forces from the main sector of the battle, where NATO would confront a concentrated attack with units from four different national armies, some of which might not be able to fight at full strength.

Compounding this problem would be a lack of theater reserves. This deficiency would rival what AFCENT has faced for many years and is the primary reason why many military authorities have feared use of the layer cake along the old inter-German border. NATO's forward corps sectors, now stretching 780 kilometers from the Baltic Sea to the Austrian border, would consume 34 DEFs. In the early stages of mobilization, NATO would have available only three DEFs as a theater reserve, which would expand to 13 DEFs once mobilization had been complete. Even so, this would amount to only about 25 percent of NATO's total posture, and NATO would still be vulnerable to a concentrated Soviet attack aimed at gaining a local breakthrough and then waging a successful war of maneuver in the rear areas. Moreover, about half of NATO's forces, especially French units and German reserves in Bavaria, would have to travel 450 kilometers cross-country to respond to an enemy breakthrough near Berlin. This distance is 150 kilometers further than what has confronted them with NATO's defense line drawn on the inter-German border. With this extra distance and the associated longer supply lines taken into account, their capacity to respond in a timely fashion would be brought even further into question.

NATO's military problems would obviously mushroom if it were to retain its linear defense and layer cake as its commitments march eastward. Impending force cuts will make the old doctrine difficult to execute on the inter-German border. The extension of NATO's defense line to the Oder-Neisse rivers would make it more problematic still, perhaps impossible.

NATO's Strategy Guidelines

In deciding what new employment doctrines to adopt during the three stages ahead, NATO will be influenced by the future strategy guidelines that were laid down by NATO's political leaders at the July 1990 London Summit. There NATO's leaders reaffirmed their commitment to the common defense and pointedly noted their intention to defend "all the territory of all of our members." They also noted that as Soviet troops leave Eastern Europe, "NATO's force structure and strategy will change fundamentally." They agreed to prepare a new military strategy that would move away from both forward defense and flexible response. The former, they said, should give way to a reduced forward presence while the latter should be revised to endorse a reduced reliance on the threat of nuclear escalation. As a result, NATO's conventional military posture would change in the direction of smaller active forces and greater emphasis on reserve formations. This reduced emphasis on high

readiness, they said, would permit reductions in NATO's training requirements and exercise programs. Meanwhile, NATO's nuclear arsenal is to be scaled back, including eventual withdrawal of short-range nuclear artillery warheads.[2]

For all their landmark qualities, these statements will require publication of NATO's new official strategy before their explicit meaning can be spelled out. They do nonetheless lay down some important guidelines, explicit and implicit, for NATO's future employment doctrine. In some respects, these guidelines create room for doctrinal innovation, but in others, they seem to circumscribe NATO's room for choice.

By relegating nuclear weapons to a "last resort" status while reaffirming the common defense, the London declaration solidified NATO's commitment to a fully adequate conventional defense posture. And by acknowledging Germany's impending unification as a full-fledged member of the Alliance, the London Summit commits NATO to extending defense protection to eastern Germany in a crisis. The need to extend NATO's employment doctrine eastward thus has received official endorsement. The London declaration made no mention of a layer cake defense. Indeed, statements by General Galvin and other NATO military authorities all point in the direction of a more flexible, mobility-oriented concept. The official groundwork thus has been laid for adoption of a better employment doctrine that compensates for the vulnerabilities of a rigid linear array along both the old inter-German border and locations eastward.[3]

At the same time, the London declaration showed sensitivity to Soviet concerns in ways that will inhibit NATO's future military flexibility in basing its forces, if not necessarily in the manner it employs them. The vague reference to a reduced emphasis on forward defense and active military forces is one indicator that the Alliance does not plan to transplant its peacetime defense posture into eastern Germany once Soviet forces have departed. This commitment to restraint was reinforced shortly thereafter at the Kohl-Gorbachev meeting in the Soviet Union. There Gorbachev publicly accepted a unified Germany's membership in NATO and evidently agreed that nationally assigned German troops could enter into eastern Germany as Soviet forces are withdrawn. But in exchange, Chancellor Kohl apparently agreed that no NATO formations, including German troops assigned to NATO, would be deployed into eastern Germany. The German-Soviet Treaty in late 1990 ratified these points.

The net effect is that even after Soviet troops have departed, eastern Germany is to be only lightly defended in peacetime. Apparently, the German government plans to deploy several thousand active troops there. But these soldiers and any reserve units will remain exclusively under German command. The bulk of Germany's 370,000-man defense establishment, including heavily armored and mechanized units, will remain in western Germany and will continue being assigned to NATO. No U.S. or other allied forces will be allowed to enter eastern Germany. As a result, the bulk of NATO's ground and air combat power will be based well to the rear of eastern Germany's exposed borders. In a crisis involving a threat to Germany, these forces could deploy forward in an effort to fulfill NATO's commitment to that nation, but they would have to move a considerable distance to take up positions near

[2]The *New York Times*, July 7, 1990, p. 5.

[3]See comments by General James Galvin on the following occasions: Address to the North Atlantic Assembly, Brussels, Belgium, February 1990; Address to the General Officer Warfighting Course, Maxwell Air Force Base, Alabama, February 1990; Remarks to the Secretary General and PermReps, Mons, Belgium, February 19, 1990; Address to the Steuben-Schurz Society, Frankfurt, Germany, January 15, 1990; Question and answer session at exercise Viking Shield, Kiel, Germany, January 25, 1990; Remarks to National Press Club, Washington, D.C., March 5, 1990; SHAPE, Public Affairs Office, Mons, Belgium.

Germany's eastern border. This basing configuration will set an important parameter for designing NATO's future employment doctrine for protecting Germany.

ALTERNATIVE EMPLOYMENT CONCEPTS FOR DEFENDING ALONG THE OLD INTER-GERMAN BORDER

In the near term, NATO will need to continue defending in the vicinity of the old inter-German border. Each of the following five alternatives would replace the present linear defense with a doctrine that is more flexible and responsive: northward shift of NATO's posture, improved counter-concentration, trading space for time, nonlinear defense, and defensive defense. To an important degree, the idea of defending along the old inter-German border already is being overtaken by events. But the five concepts to be reviewed here, all of which emerged in Western military thinking in the late 1980s, have a more enduring quality regardless of their specific geographic focus. Since they can help form the building blocks for an employment doctrine aimed at defending eastern Germany, they merit a careful review.

The analysis assumes that NATO's forces will be cut by about 10 percent over the next two years (rather than a 25 percent cut over five years) and that the Soviets would be capable of mobilizing about 50–65 divisions (DEFs) by M+30. These assumptions produce the ground balance displayed in the Table 7. NATO initially would find itself outnumbered by about 1.25–1.6/1, and even a 10 percent cut would have thinned its reserves. The main risk is that the Soviets might concentrate opposite NORTHAG and overpower NATO's defenses there. The defense challenge facing NATO would be to ensure that the initial disadvantage did not prove fatal. Alterations in NATO's theater employment doctrine would be designed to help achieve this demanding goal.

Northward Shift of NATO's Posture

Shifting NATO's force posture northward would place greater emphasis on NORTHAG at CENTAG's expense. This concept takes advantage of the complete withdrawal of Soviet forces from Czechoslovakia by 1991 along with that nation's defection from the Warsaw Pact. As a result, any Soviet attack against NATO is likely to be launched almost exclusively across the North German Plain. This northward shift in the Soviet threat causes NATO to transfer its center of gravity in that direction, strengthening NATO's defenses in the area most likely to be attacked and reducing the likelihood that NATO's forces could be caught out of position.

Table 7

NATO/Soviet Buildup Rates (DEFs in 1991)

Day	Soviets	Nato	Ratio
M+30	50–65	40	1.2–1.6:1
M+45	50–65	45	1.1–1.4:1
M+60	50–65	49	1.0–1.3:1
M+75	50–65	52	1.0–1.25:1

The two-step process implementing this change involves first taking NATO's reductions exclusively from forces now allocated to CENTAG, thereby sparing NORTHAG entirely. Table 8 shows how this preferential reduction would help alleviate the risk that NORTHAG would be denuded against a Soviet military threat. NORTHAG would now have the 53 percent of NATO's forces that CENTAG had before. More important, NORTHAG's present posture would still be able to form a strong front line and, counting U.S. III Corps and other Allied forces, a small reserve.

The second step would involve the thinning out of CENTAG's remaining forces to transfer more units to NORTHAG. In this scheme, U.S. V Corps would be withdrawn from its present frontline role of defending in the Fulda Gap sector. It would be withheld as an AFCENT reserve for likely assignment to NORTHAG, where it would join U.S. III Corps. In addition, CENTAG's reserves would be reduced by also shifting 2.0 DEFs of the French First Army to NORTHAG. Table 9 shows how this readjustment would affect NATO's posture. NORTHAG's posture would be bolstered by 5.6 DEFs, a 25 percent gain. NORTHAG's reserves would increase 100 percent. CENTAG meanwhile would be left with a reduced posture of 13.2 DEFs.

The Soviets could take advantage of CENTAG's reduced posture by committing forces in an effort to gain a breakthrough in the Fulda Gap or elsewhere north of the Czech border. NATO presumably would compensate for the withdrawal of U.S. V Corps from this area by extending the frontages of the three corps remaining in CENTAG (U.S. VII and German III and II). These forces, however, would be left strung out on frontages averaging 45 kilometers, with only about three DEFs of the French First Army providing reserve support. Unless rectified, this situation could create a vulnerability of its own.

As Fig. 11 shows, extending NATO's defense line about 25–50 kilometers into eastern Germany in the area between Saltzgitter and the Czech border could rectify this problem. This step would make the Kassel-Fulda-Meinengen region less vulnerable and straighten the concave bulge in NATO's center that has long been a major source of worry. NATO's overall defense line would be reduced by 100 kilometers. The CENTAG forces still deployed in this area (about seven DEFs) would have their divisional frontages reduced from about 40

Table 8

Effect of Preferential NATO Force Reductions (DEFs at M+30)

	Present	Symmetrical Reductions	Preferential Reductions
NORTHAG			
Forward corps	16.0	14.2	16.0
Reserves[a]	5.2	4.6	5.2
Subtotal	21.2	18.8	21.2
CENTAG			
Forward corps	17.5	15.6	13.9
Reserves	6.3	5.6	4.9
Subtotal	23.8	21.2	18.8
AFCENT Total	45.0	40.0	40.0

[a]Assumes U.S. III Corps and France's III Corps and FFAR are counted as NORTHAG reserves.

Table 9

Effect of Transferring Forces to NORTHAG (DEFs at M+30)

	Preferential Reductions	NORTHAG-Oriented Employment Doctrine
NORTHAG		
Forward corps	16.0	16.0
Reserves	5.2	10.8
Subtotal	21.2	26.8
CENTAG		
Forward corps	13.9	10.3
Reserves	4.9	2.9
Subtotal	18.8	13.2

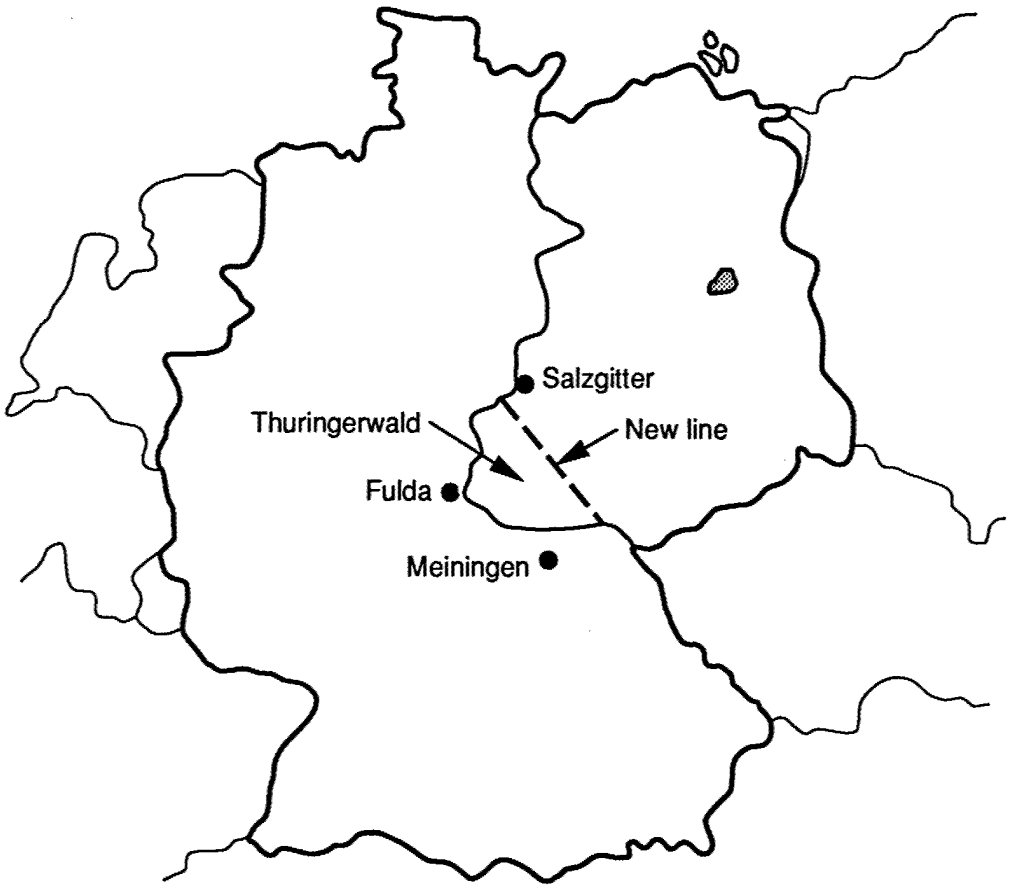

Fig. 11—Effect of Altering NATO's Defense Line

kilometers apiece to about 30 kilometers, a now manageable distance. Equally important, these forces would now be able to anchor their defenses on the Thuringerwald forest, which would provide a formidable barrier against a Soviet advance. The net effect of extending NATO's wartime defenses into eastern Germany in this small way would be to render CENTAG secure, thereby allowing for a sizable diversion of forces to NORTHAG.

Provided these two steps were feasible, the effect on NORTHAG's defenses would be salutary. The local imbalance there would be reduced from 2–3/1 to a more manageable 1.5–2/1. Commitment of CENTAG's remaining reserves would reduce it further to 1.35–1.85/1. Equally important, NORTHAG's posture now would be able both to erect a strong front wall and to form a large 10-division reserve. The process of reorienting CENTAG forces to NORTHAG would involve important changes in NATO's defense plans and logistic support system. But these changes would enhance NORTHAG's ability to quickly marshal large reserves to control an enemy breakthrough. Table 10 shows the effects on NORTHAG's time-phased buildup rate.

NATO's air strategy would be similarly affected. NATO's plans for performing its air missions would focus on the NORTHAG battle in much the same way that the ground defenses would be reoriented. Implied here is a growth of 2 ATAF (NATO's numbered air force in NORTHAG) at the expense of CENTAG's 4 ATAF. The past defense situation called for a roughly equal distribution of NATO's aircraft and sorties between these two commands, but the new environment would give 4 ATAF preference. Since NATO's air doctrine has long called for the flexibility to swing back and forth between both commands, the change required would probably not be as disruptive as that inflicted on NATO's ground doctrine, but it still would be an important one.

In the years immediately ahead, NATO is likely to enjoy a small numerical preponderance in tactical air forces. Assuming U.S. reinforcements have fully arrived by M+30, about 3250 NATO combat aircraft would face roughly 2500–3000 Soviet aircraft. This numerical advantage, coupled with their qualitative edge, would place NATO's air forces in a better position to play the "ace in the hole" role that Western strategists have long hoped for them.

NATO's efforts to prevent a Soviet breakthrough would still depend upon the capacity of its air forces to perform close air support and battlefield interdiction missions. Complicating

Table 10

Effect on NORTHAG's Buildup Rate in Main Sector (DEFs)

	No Northward Shift	After Northward Shift
D-Day	2.6	3.0
D+2–6	7.2	13.8
D+7–10	12.8	18.0
D+11–14	16.4	18.0

NATO's efforts to focus its airpower on the ground battle would be the continued need to perform the air defense and counterair missions. Since the overall enemy air threat would be smaller than in years past, the demands on NATO's air forces for these missions would be

reduced somewhat. However, the growing likelihood that NATO will rely more heavily on plans to move its ground forces across the battlefield would increase the premium on controlling the battlefield airspace. Although a small number of enemy air strikes would cause little damage, they could seriously disrupt and slow NATO's movements in its own rear areas. Especially in the early stages, screening NATO's rear areas and suppressing enemy airbases could consume several hundred scarce NATO sorties daily.

Although the deep interdiction mission would still be necessary, it might consume fewer sorties than before. The principal reason is that the Soviets, by deploying a ground posture of only 50–65 divisions, would present a considerably smaller second strategic echelon than before. The 10–20 divisions still employed as a second echelon would play an important role in Soviet strategy and would still require targeting. But because of its reduced size, the second echelon would absorb fewer air sorties than before, which could free some NATO aircraft for close air support and battlefield interdiction missions.

Determining how these competing demands would affect NATO's air strategy lies beyond the scope of this study. A new air campaign plan that achieves an appropriate balance among NATO's multiple missions would have to be devised. Regardless of what level of air support can be applied to the ground battle, a strategy that concentrates it on NORTHAG would make sense.

For both ground and air forces, a NATO effort to shift emphasis toward NORTHAG necessitates a U.S. willingness to accept a further change in broadening its geographic scope of responsibility in Central Europe. It also would require a NATO effort to learn how to absorb and employ more U.S. forces in NORTHAG. The steps that already have been taken, over the past decade, to employ III Corps and some five USAF tactical air wings in NORTHAG has enabled NATO to initially develop this capability. In the employment doctrine envisioned here, at least one additional U.S. corps would deploy to NORTHAG in a crisis. Equally important, several additional USAF wings would be allocated there. NATO's success in bolstering NORTHAG would depend heavily on its capacity to develop both operational plans and a logistic support capacity for employing these forces to best advantage.

What applies to U.S. forces also holds true for French forces. In recent years, both the French III Corps and Rapid Reaction Force have begun to develop a NORTHAG orientation. Contemplated here is the wartime deployment of an additional French corps from CENTAG along with French air support. This step clearly could be undertaken only if France is willing to move one step further in the direction of returning to NATO's integrated command. NATO would then also face the task of designing appropriate missions for French forces to perform in NORTHAG.

On the surface, shifting NATO's defenses toward NORTHAG offers attractive military benefits. But it also carries with it some costs that must be factored into the equation. The act of shifting NATO's military posture would bring about potentially unsettling changes in NATO's internal political balance. It would require U.S. and French willingness to accept changes in their traditional political commitments to NATO. Neither nation's acquiescence can be taken for granted, especially France's. If either or both of these nations proved unwilling, NATO could still pursue this alternative by reorienting Germany's III and II corps northward. This step, however, would leave the defense of southern Germany largely in the hands of the United States and France, a development that might unnerve the German government. In any event, Britain's preeminent role in NORTHAG, and its influence in NATO circles, might diminish. London might not welcome such a trend. Perhaps more important, Belgium, the Netherlands, and Denmark might take advantage of this change to

scale back their contributions to NATO, causing a loss of some of the gains of this alternative.

This alternative also would bring other costs in its wake. Major changes in NATO's military plans always have a disruptive, stressful effect on the Alliance's civilian and military staffs. This case would be no exception, and it would occur when larger matters are straining NATO's cohesion. Additionally, some budgetary costs would ensue. For example, U.S. and French (or German) ammunition and war reserve stockpiles presumably would have to be moved northward. Facilities would have to be found and scarce NATO infrastructure funds committed to building them. NATO's training and exercise profile also would have to change, at a time when field training exercises are to be cut back. Failure to implement these programmatic measures would take the edge off this alternative's effect.

If these political and budgetary barriers could be surmounted, this alternative would provide greater leverage in solving NORTHAG's vulnerability than any of the other alternatives. Even if full implementation proves infeasible, partial steps (e.g., revised plans with no programmatic changes) still would be helpful. This alternative makes military sense since the Soviet threat to southern Germany is rapidly disappearing while the northern region remains vulnerable. It also could have a useful influence on impending budgetary debates in NATO's capitals. If they remain anchored in CENTAG, a large portion of NATO's conventional forces will find themselves lacking a sufficiently prominent mission to justify the large sums of money spent on them. By pointing them northward, NATO would restore them to a central role in Alliance defense strategy while helping meet security requirements.

Improved Counter-Concentration

By shifting its center of gravity northward, NATO would strengthen its ability to mass large forces in the critical NORTHAG area, but NATO could still find itself facing a difficult task, especially if the Soviets concentrated to the fullest. In that case, the Soviets, with an initial edge of 1.5–2/1, would still hold the kind of advantage that could produce a decisive victory in the event of a major breakthrough.

Moreover, the Soviets would have enough forces there—40–55 divisions—to exert strong offensive pressure against vulnerable areas of NORTHAG's line. These areas include, in particular, the Luneburger Heide in the Dutch corps sector, the Braunschweig area guarded by British forces, and the Gottingen Gap defended by Belgian forces. Simultaneous Soviet attacks against all three areas would complicate NATO's efforts to keep the situation under control. The outcome could depend heavily on NATO's ability to respond to the ebb and flow of battle by maneuvering its forces tactically on the NORTHAG battlefield. To the extent NATO proved quickly able to shift forces to contain local penetrations, the likelihood of a successful effort would increase. To the extent NATO failed, the risk of a breakthrough and eventual defeat would rise.

The second doctrinal initiative to be examined here is designed to facilitate NATO's capacity to perform this important mission by enhancing its ability to rapidly redeploy forces across the NORTHAG battlefield. It would preserve a strong NATO frontal wall in NORTHAG, but it also recognizes that the attacker's advantages of surprise, the initiative, and concentration are likely to produce a breakthrough. A strong capacity to counter-concentrate in a timely fashion is therefore needed. In calling for this capability, this initiative takes its historical inspiration from the redeployment operations that the U.S. Army conducted at the Battle of the Bulge (see Section 4).

This alternative calls on NATO to develop the capacity to withdraw forces from less threatened sectors in NORTHAG and quickly deploy them to any point where an enemy breakthrough appears imminent. It also calls for NATO to develop the ability to conduct operational fires on a cross-corps basis. Additional ground combat forces would be needed to increase NATO's local force-to-space ratio, to restore a cohesive line along expanding salients, and to conduct flanking attacks into advancing enemy formations. Firepower provided by such indirect systems as long-range artillery and ATACMS missiles could help NATO's commanders restore the situation by inflicting high attrition on enemy forces, by disrupting their capacity to attack in a synchronized and coordinated fashion, and by delaying their advance.

An attractive feature of this initiative is that it could, in theory, be pursued inexpensively. Nearly all of NATO's ground combat forces are already mechanized and therefore sufficiently mobile to achieve high cross-country mobility rates, especially U.S., German, and British forces. This doctrine thus would not require major increases in U.S. and allied programs for buying new weapon systems and otherwise modernizing NATO's ground and air combat forces. It is compatible with the era of reduced defense expenditures ahead.

What this initiative would require is accelerated coalition planning in many areas. As matters now stand, NATO would suffer from the liability of trying to fight a more mobile war as a multinational coalition with forces drawn from seven nations. NATO's defense plans historically were designed primarily on a corps-by-corps basis, with less attention paid to the task of forging these separate corps into a cohesive whole. The Bagnall reforms of the mid-1980s, reinforced by the introduction of AirLand Battle and FOFA, helped strengthen NATO in this area, but further improvements would be needed to enhance NORTHAG's capacity to act as an integrated Army Group. In endeavoring to strengthen NORTHAG's agility in counter-concentrating, this doctrine takes aim at bolstering NATO's capacity to wage a coalition war in a maneuver battle where timely responses and synchronized operations would be at a premium.

A key to this doctrine would be an improved capacity to move more ground combat forces in a shorter time than before. In the past, rapid redeployment operations have largely been the province of AFCENT reserve forces, especially U.S. III Corps and a few German and other Allied units. Other units have focused on maneuvering within their individual corps sectors. This initiative would dramatically expand the scope of NORTHAG's cross-corps movement activities. If previously only about 25 percent of NORTHAG units would have been required to conduct cross-corps movements, now perhaps 75 percent would need to become skilled in this mission.

Such a change would require a major alteration in NATO's practices and doctrinal thinking. Plans would need to be drawn up to permit coordinated movements on a larger scale than previously envisioned. NATO's ground forces would need to focus a larger portion of their training on the task of disengaging rapidly from their positions and laterally moving long distances elsewhere. Perhaps most important, NATO would need to harmonize its tactical doctrine to a greater degree than previously attempted. This step would be vitally important to any effort aimed at conducting coordinated maneuver operations with forces drawn from separate parent corps and entirely different nations.

Although this doctrine could be implemented with highly mobile ground forces, its programmatic thrust lies in the direction of acquiring larger helicopter forces for attacking enemy armored forces and transporting NATO's troops. This trend already is evident in the U.S. Army and the French FFAR. Evidently the German Army now plans to transform three

divisions into "air mechanized units" that would rely heavily on helicopters for their movement and antiarmor firepower. The advantage of helicopters lies in their high mobility and in their capacity to attack enemy tanks from the top, where their armor is weakest. Their disadvantage lies in their high acquisition and maintenance costs and in their comparative fragility. Although tests and exercises suggest that they will be survivable against the lethal Soviet air defenses of the modern battlefield, this conclusion has yet to be tried in combat against a well-armed opponent.

A predominantly helicopter-equipped force also would suffer from operational liabilities. Transport helicopters can lift infantry troops to the vicinity of battle, but they often cannot carry them into close engagements. Infantry fighting vehicles are needed for this purpose. And although attack helicopters can deliver defensive antiarmor firepower, they are less useful as offensive instruments, especially for delivering shock action. Tanks are still needed for that function. These considerations do not argue against the acquisition of additional helicopters, but they do argue for retaining the balanced mix offered by combined arms forces.

The original idea behind ATACMS was that this long-range missile would be fired largely on an east-west axis at second echelon targets well behind enemy lines. This mission would not disappear entirely, but modified plans would have to be developed for firing ATACMS on a north-south axis, and possibly against enemy forces that were in the process of penetrating NATO's defense lines. ATACMS itself is capable of firing in any direction, but NATO's C3I system would need to become attuned to the task of discriminating between friendly forces and enemy targets. This task, requiring real-time coordination between local ground commanders and firing units miles away, could prove highly demanding. Among other things, it would require advance practice in training exercises.

NATO's plans for employing its close air support operations also would have to be fine tuned in similar ways. In essence, NATO's air operations would have to be closely integrated into its ground scheme of maneuver in controlling what is likely to be an increasingly open, fluid, and chaotic battlefield. U.S. forces have long emphasized close air-ground coordination, but this is less true for allied forces. Allied efforts to improve in these areas would need to accompany a NATO doctrine of enhanced counter-concentration.

Adoption of this doctrine would require NATO to take strides toward developing an improved NORTHAG logistic system. The present system of making logistic support a national responsibility could seriously retard NATO's ability to shift combat forces from one corps sector to the next. Ideally, creation of a fully integrated NATO multinational logistics system, entailing establishment of a NATO support command, would be desirable. Short of this step, measures would need to be taken to make NATO's logistic structure interoperable. That is, each nation's logistic forces would need to become capable of providing support to units from other nations at least on an interim basis.

Some steps have already been taken in this direction. A good example is the long-term defense plan of the 1970s which sought to foster improvements through such measures as procurement of common ammunition, the collocated operating bases program, and Stage B cross-servicing. In the 1980s, the conventional defense improvement effort sought to achieve better logistic integration through enhanced armaments cooperation and common weapon procurement. Nonetheless, much remains to be done, especially for NATO's ground forces. Particularly important to a counter-concentration doctrine would be development of a better capacity to share artillery ammunition, petroleum, and other stocks on a multinational basis. Also important would be an enhanced capacity within each corps sector to provide, on a surge basis, transport of ammunition, maintenance and engineering support, and military police

for traffic control. Without measures in these areas, any effort to develop a redeployment doctrine might fall short of aspirations even if the necessary steps were taken to prepare NATO's combat forces for their new roles.

In summary, this doctrinal alternative, which evidently has taken hold in German military circles, offers a prescriptive solution for curing some of NATO's defense ills arising from the thinned-out modern battlefield. Its advantage is that it promises tangible returns at low financial cost. The cost might be higher than appears at first glance, however, if a major shift to helicopters and altered logistic support forces proves necessary. Wihtout a budgetary spiral, its principal disadvantage lies in the added demands that would be placed on NATO's coalition planning at a time when political resolve is slackening.

These strains would be least onerous if this doctrine were adopted within the framework of existing national corps. They would be greatest if the doctrine were applied on a theaterwide basis in an effort to achieve greater integration of these corps. It is precisely in the latter category, however, where the greatest innovations would be needed. NATO's main problem is not that its individual corps cannot fight effectively, but rather that its separate corps might not be able to fight together as a cohesive whole. This doctrine points toward a solution to this problem that clearly merits pursuit. But the departures that it requires, which involve finding solutions to problems that have badgered NATO for over two decades, are far from trivial.

Trading Space for Time

The first two alternatives would strive to preserve NATO's forward defense concept, along with a strong frontal wall, by introducing greater lateral mobility into NATO's posture. The third alternative would also preserve a frontal wall, but it would make this wall less brittle by relaxing the commitment to protect all of Germany's territory. It would not abandon linear defense per se, but it would alter its positional quality. In essence, it would reintroduce the concept of tactical retrogrades into NATO's defense plans by allowing for localized retreats when necessary. The task of recovering lost ground would be deferred until after the enemy advance has been halted and a counterattack becomes possible. This alternative thus views the forward defense concept in flexible rather than literal terms and subordinates it to the more important goal of conducting a sound defense.

The concept of "trading space for time" is deeply ingrained in military vocabulary, but it is nevertheless something of a misnomer. This doctrine does trade space, but it does not aspire to buy time as a goal in itself: Time is meaningless if it merely delays inevitable defeat. This doctrine trades space and buys time to enable the defender to slip the attacker's first punch and pull himself together. Applied to Central Europe, its goal would be to reduce a serious vulnerability in NATO's posture, one that would become worse as its forces are reduced. This doctrine anchors itself on the proposition that any effort to cling stubbornly to every inch of territory can contribute heavily to the defender's undoing. A good historical example is Hitler's refusal to disengage from Stalingrad, an exercise in futility that cost the Wehrmacht dearly when von Paulus' 6th Army was trapped by a Russian double envelopment attack. By rigidly maintaining its forward positions, this doctrine reasons, NATO could expose itself to similar risks, especially if reserves are not immediately available.

In theory, two problems could arise if NATO's defense effort was conducted in a rigidly positional, territory-holding way. First, NATO units in the Soviet main sector could become vulnerable to being either encircled and cut off in the face of local penetrations or physically

destroyed as a product of intense fighting and high attrition. As a result, the risk would mount that an otherwise benign enemy advance might be transformed into a full-scale breakthrough. Second, by remaining in their positions despite this unravelling, NATO's forces in adjoining sectors could themselves become exposed to a subsequent enemy encirclement. This doctrine proposes to solve both problems by allowing NATO's forces in NORTHAG to withdraw in the face of an advance, thereby enabling them to constantly slip the Soviets' punch. The Soviets might conquer some territory, but they would not succeed in fracturing NATO's cohesion. NATO would be able to continue fighting, and the time afforded by retreating would enable it to redeploy reserves to the main sector.

By allowing withdrawals, this doctrine also would endeavor to transform the battle's attrition dynamics by reducing NATO's casualty rate. The risk would be reduced that some NATO units might be captured. Equally important, NATO's forces in theory would be better able to extract favorable exchange rates. In essence, they would be able to fire at exposed enemy forces from a safe distance and behind prepared positions. This practice would enable them to inflict high losses on the enemy, while always withdrawing once enemy forces drew near. Consequently, NATO's forces would avoid fighting the close battles that an attacker normally requires to destroy the defender.

The result presumably would be an attrition process highly favorable to NATO. A good historical example is the yielding defense that outnumbered Israeli forces conducted against the Syrian Army on the Golan Heights in 1973. In this case, the Israelis extracted favorable exchange ratios of about 5–7:1 or more. By the time the Syrians had advanced through the Golan Heights, they had been badly depleted. This left them vulnerable to an Israeli counterattack once Israeli armored reserves had arrived on the scene. The same outcome could be orchestrated in Central Europe. By inflicting heavy casualties on the Soviets while buying time for NATO's reserves to arrive, NATO could come from behind to win. Although German territory would be sacrificed in the initial stages, the theory holds, NATO eventually would gain the numerical ascendancy and then be able to counterattack decisively to regain the lost ground.[4]

Adoption of this doctrine would not require abandoning the practice of establishing a frontal array, but it would entail major revision of the employment concept that NATO has followed since the mid-1970s. This concept placed NATO's general defense positions (GDPs) near the inter-German border and called upon NATO's forces to make a determined stand there. Implementation of this doctrine would begin in peacetime with a rearward movement of particularly exposed GDPs to establish them on better ground. In an actual war, NATO's defense effort would begin with covering force operations by armored cavalry troops aimed at identifying the enemy's main effort and channeling it into NATO fire zones. NATO's front-line units would begin their resistance efforts at their forward GDPs, but rather than fight a pitched battle there, they would begin withdrawing from one phase line to the next. As they withdrew, they would continuously bathe the enemy in lethal direct and indirect fires. Along the way, NATO's forces might conduct limited counterattacks when the conditions were right, but NATO's principal tactics would be defensive. This process would continue until either the enemy attack had ground to a halt or NATO's troops had arrived at advantageous terrain features where a positional stance made sense, or enough reserves had arrived to thicken NATO's line sufficiently.

[4]For an insightful quantitative analysis of how a defender can use retrograde tactics to achieve favorable exchange rates, see Epstein, 1990.

In theory, NATO's withdrawal tactics might be confined to the enemy's main attack sector, which might cover about 100 kilometers of the front line. An effort to confine withdrawals to this sector while holding forward elsewhere, however, could lead to the development of the kind of dangerous bulge that emerged in the Ardennes in 1944. This situation would create strains of its own by stretching the frontages of NATO's committed units, perhaps beyond manageable limits. For the purpose of lessening this risk, NATO's withdrawals might be extended well beyond the main sector to maintain a continuously smooth front line. In the extreme case, NORTHAG's entire front wall might fall back in tandem, resulting in the slow but steady movement of NATO's defense line into western Germany.

The amount of German territory that might be lost temporarily would depend upon a complex set of battlefield dynamics that cannot be forecast in advance. The most likely development would be that NORTHAG's defense line would pivot on its southern hinge and swing backward like an opening door, with the greatest withdrawals taking place in the far north, near Hamburg. Assuming the Soviets kept pressing their advance, a natural stabilization point would be reached with a NATO defense line running on a northwesterly axis along the Leine or Weser rivers, which form the western boundary of several NORTHAG corps. In other words, NATO's forces would have yielded some 30–50 kilometers along NORTHAG's southern boundary, roughly 100 kilometers in the middle sectors, and about 150 kilometers of territory in the far north.

A Soviet advance might grind to a halt, of course, well east of the Weser River. A possible NATO withdrawal this far is suggested by the likelihood that a concentrated Soviet attack could continue its forward drive long enough (1–2 weeks) to achieve this distance before exhausting itself. Also, the NORTHAG terrain could compel NATO to make a withdrawal of this magnitude. In southern NORTHAG, the terrain favors a stiff defense effort, since the Solling and Harz mountain ranges would bar the Soviets' way and the Weser River lies only 25–50 kilometers to the rear. Farther north, the terrain opens and favors the attacker. A natural defense barrier lies along the Elbe-Sieten-Mittelland Canals near the old inter-German border. West of these canals, there is open terrain and, urban areas aside, no natural barriers until the Aller-Leine-Weser river complex is reached. NATO's forces would therefore be likely to continue fighting a yielding mobile defense until they reached these rivers. Since further withdrawal would allow the Soviets to gain access to open terrain and the lowlands, NATO's forces most likely would establish firm positions on the Weser River.

Although a serious setback, a withdrawal as far west as the Weser River would buy NATO valuable time in which to move forces from CENTAG northward and to bring additional reinforcements from the United States. NATO's prospects for a successful counterattack at this juncture would depend heavily on how many losses the Soviets had incurred in reaching the Weser River. Table 11 illustrates some of the sensitivities by displaying postattack force ratios (NATO vs. Soviet) as a function of exchange rates, the size of the Soviet attack, and NATO reinforcement from the United States after D-Day. Provided NATO achieved favorable exchange ratios, it would be left with an overwhelming advantage against a Soviet invasion force of 50 divisions or less. Against a larger enemy force, any NATO superiority would depend upon the arrival of U.S. reinforcements. In the best case, however, an advantage of 1.8:1 should allow NATO to successfully counterattack against a Soviet posture of only 20 DEs spread quite thin.

This employment doctrine offers one militarily feasible way to compensate for the risks of a NATO linear defense at somewhat lower force levels than now. Above all, it could accomplish its minimal objective of preserving NATO's cohesion and thus deny the Soviets an

Table 11

NATO/Soviet Force Ratios in NORTHAG (DEFs at Weser River)

	Soviet Invading Force (DEFs)			
	50		65	
Exchange ratio favoring NATO	3–5:1	6–8:1	3–5:1	6–8:1
NORTHAG force ratio				
NATO D-Day forces	1.0/1	4.0/1	.7/1	1.2/1
Fully reinforced	1.5/1	7.0/1	1.0/1	1.8/1

opportunity to gain a decisive military victory. By insuring that the Soviets at best would pay a high price for gaining a small amount of territory, it would help preserve a still-strong conventional deterrent for NATO. Beyond this, it might work well enough for NATO's forces to win the attrition battle to such an extent that they later could swing over to the offensive and seek either to regain lost ground or to gain an operational victory of their own. This added element of uncertainty facing the Soviets would further strengthen the psychological effect of NATO's conventional deterrent and defense posture.

In the absence of a prompt capacity to redeploy forces from CENTAG, the risks created by a thinned-out NORTHAG posture might leave NATO with little alternative but to implement this doctrine at least to some degree. However, the doctrine of trading space for time does have military downsides. In particular, it might be less easy to implement than surface appearances suggest. A widespread retrograde operation can be among the most difficult to carry off. If done inexpertly, it can create enemy breakthroughs of its own. A retreat, in this case, easily could deteriorate into a rout. Even if a phased withdrawal is successfully executed, the goal of conducting a later counterattack would depend heavily on NATO's ability to inflict enough losses on the Soviets to leave their forces too spent to resist even fatigued NATO forces. The Israelis achieved this goal, but against an inferior army. The same might not be the case against the Soviet army.

An equally important drawback is the likelihood that this doctrine might leave some German territory more vulnerable to seizure than would a firm frontal defense. In particular, the Soviets might be able to advance some distance, then halt and sue for peace from a military position that could not readily be dislodged. Because this doctrine makes withdrawals contingent upon military exigencies and subject to later rectification, it is not inconsistent with the goal of forward defense. But neither is it the kind of doctrine that is likely to inspire strong Alliance unity under the banner of protecting all of NATO's territory, as the Washington Treaty demands.

Nonlinear Defense

This employment doctrine contrasts sharply with the first three alternatives by calling for a major departure from linear defense, positional or otherwise. The underlying concept of the first three doctrines was to preserve NATO's frontal wall by taking steps to shore it up. Two of these doctrines provided schemes for bolstering the front wall by redeploying reserve

forces along its breadth; the third endorsed phased withdrawals to buy time for this purpose. The "nonlinear" defense doctrine departs almost entirely from the philosophy of creating a strong front line that consumes a large portion of NATO's ground forces. It calls instead for NATO's center of gravity to be shifted well to the rear, with only a thin line of troops deployed forward. Beyond this, it promises to replace a linear array's reactive mentality with a very different style of military thinking.

This doctrine endeavors to turn the maneuver warfare tables on the Soviets by having NATO adopt its own brand of defensive mobile warfare. In viewing the battlefield in far more fluid terms, it would pursue NATO's defense objectives principally through a combination of long-range firepower and mobile counterattacks. By any standard, this would be a dramatic change in how NATO implements its conventional defense strategy

The nonlinear defense doctrine has found favor with the U.S. Army, which is endeavoring to update it to modern conditions, but its origins (in somewhat different form) lie with the Germany Army's experience on the Russian front in World War II. There the Wehrmacht found that fixed positions and linear arrays regularly crumpled in the face of powerful attacks by mobile mechanized Soviet forces. Mobile warfare, relying on forward strongpoints but with counterattacking armored reserves carrying the main load, worked far better. The Russians ultimately won that war because of their vastly superior numbers, but they were never able to deliver the knockout punch that they constantly sought.[5]

The nonlinear doctrine is skeptical of the notion that a thick line of prepared positions gives the defender major advantages. Rather, it reasons, the attacker probably would be able to employ his own advantages in surprise, initiative, concentration, and freedom to select the time of attack to overpower the defense. This especially is the case in the early stage of a war, which is a unique phase because defending troops have not yet had an opportunity to gain combat experience. Consequently, it holds, the Soviets could employ artillery and air to batter NATO's positions and then finish them off with infantry and armor, with a high risk that NATO's frontal defense would be punctured and then unhinged. By continuing to hamstring itself with a linear defense array that makes even less sense at lower force levels, NATO presumably would help tilt the warfighting balance in the Soviets' favor.

The nonlinear doctrine calculates that a positional defense would sacrifice many of the advantages in leadership, command style, combat spirit, flexibility, and operational skill that Western forces allegedly possess over their Soviet counterparts. Above all, it does not believe that a defensive mentality and head-to-head clashes are the best way to defeat an attacking force. It places greater stock on the tactical offensive, and it prefers to launch flanking attacks and bold assaults against the enemy's own rear areas. It is thus heavily anchored on mobile operations that strike at the enemy's weaknesses rather than at his strengths. It would seek to beat the Soviets at their own game.

The extent to which this nonlinear doctrine would alter NATO's employment concept is illustrated in Table 12, which displays a spectrum of alternatives for deploying the forces of a NATO forward corps of four divisions assigned to a 90 kilometer sector, beginning with a pure linear array and culminating with a nonlinear defense. In the linear model, all four NATO divisions are deployed directly on the front line, with frontages averaging about 22 kilometers. Its defensive philosophy would be entirely positional with little capacity to redeploy elsewhere. The AirLand Battle model represents NATO's current employment doctrine, with three divisions forward and one division held in tactical reserve about 25 kilometers to

[5]See von Mellenthin, 1979.

Table 12

Employment Alternatives for a NATO Forward Corps

Divisons	Linear Defense	AirLand Battle	Strongpoint Defense	Nonlinear Defense
Front-line	4	3	2	1
Shallow reserve		1	1	0
Deep reserve			1	3

the rear. It would rely on a combination of forward positions supplemented by limited counterstrikes.[6]

The "strongpoint" model deploys two units forward on 45 kilometer frontages, one in tactical reserve, and one in deep reserve about 50–100 kilometers to the rear. Its frontal forces would not be densely deployed enough to create a contiguous front line, but they could establish local strongpoints. This model would rely more heavily on counterattacks to defend its sector in strength, but it also would be postured to conduct a phased withdrawal while quickly deploying one or two divisions to other sectors. By comparison, the nonlinear model would leave at most one division thinly deployed along the 90 kilometer front. The remaining three divisions would be held in reserve. Moreover, all three divisions would be held well back of the front line—about 100 kilometers—in what amounts to an enclave posture. This corps would employ its combat forces almost exclusively in mobile operations, and it would be ideally situated for redeploying large forces elsewhere.

Like the Wehrmacht's doctrine in World War II, the nolinear doctrine concentrates on seizing the initiative, forcing the pace of battle, and dominating the enemy psychologically. It places a similar faith in rapid actions, a fast tempo, high mobility, great agility, and bold actions. But it would not seek to recreate the same mobile warfare tactics that the Wehrmacht employed on the Russian front. Indeed its employment scheme relies on modern technology in ways that make it quite different from what the Germans sought to accomplish.[7]

The Wehrmacht lacked modern intelligence collection systems and therefore did not have a clear picture at all times of the battlefield or enemy deployments. Nor did it have, apart from the Luftwaffe, the benefit of long-range fires. Faced with uncertainty and lacking the ability to target enemy forces at long distances, it relied heavily on maintaining forward strongpoints in fixed positions on key terrain features. These strongpoints would remain at their locations even if bypassed to exert control over the shoulders of the enemy's advance. With the enemy's forces thereby fixed and channeled into predictable directions, Whermacht armored counterthrusts would then strike them on their flanks and either destroy them or drive them back to their original departure point. In this way, the forward strongpoints would be rescued and the original German line restored.

By contrast, the nonlinear doctrine would largely dispense with forward strongpoints, relying instead on improved intelligence systems to perform many of their functions. Also, it would not rely as heavily on counterattacks as its principal warfighting mechanism. Instead, it would emphasize a combination of operational fires and maneuvers to destroy the enemy.

[6]For a treatment of the AirLand Battle doctrine, see U.S. Army, 1986. The analysis of nonlinear defense draws on U.S. Army Training and Doctrine Command, 1990.

[7]For a description of the German strongpoint model on the eastern front, see von Mellenthin, 1984.

These concepts can best be understood by an examination of how the nonlinear doctrine views the modern battlefield. Operationally, it would divide the battlefield into three parts: a detection zone extending deep into enemy territory, a battle zone reaching about 100 kilometers into friendly territory, and a dispersal area where NATO's ground combat forces would be clustered. Figure 12 illustrates these zones.

The combat process would begin as enemy forces pass through the detection zone. There they would be tracked by NATO ground, air, and space sensors that would provide intelligence data on their locations and direction. When they enter the battle zone, they would encounter NATO cavalry troops and thinly deployed combat formations. Fighting would begin at this juncture, but the main function of NATO's forces would be to conduct reconnaissance and sculpture the battlefield. They would provide additional intelligence data on the enemy forces and attempt to channel them in predetermined directions. Activity by NATO air reconnaissance units and other collection platforms would also pick up.

Shortly thereafter, NATO's forces, benefitting from real-time information flowing into intelligence fusion centers, would commence firing with tactical air strikes, attack helicopters, and long-range missiles (e.g., ATACMS). From that point forward, enemy forces would be subjected to a withering barrage of NATO operational fires aimed at inflicting high casualties and causing severe disruption. The bulk of NATO's ground maneuver units initially would be withheld from battle in their rear enclaves. As enemy forces advanced through this cauldron of fire deeper into the battle zone, NATO's maneuver units would march forth from

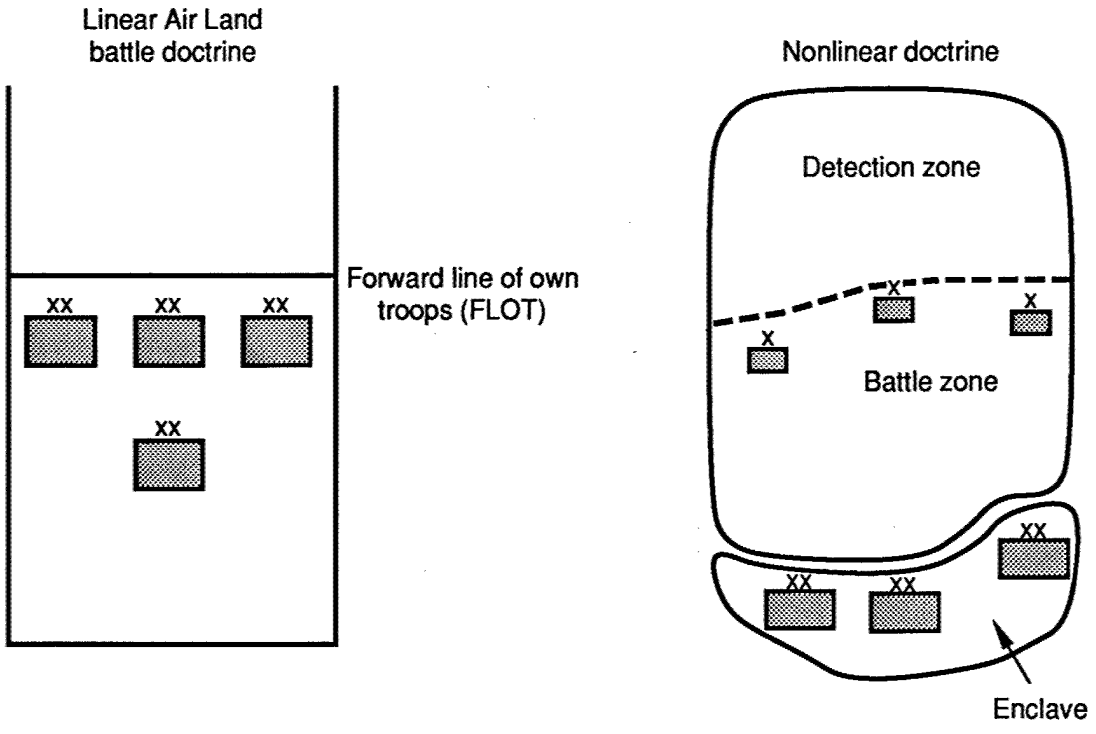

Fig. 12—Linear and Nonlinear Battlefields Compared

their enclaves and conduct offensive strikes against them. NATO armored formations would slash into their flanks and drive into their tactical rear in search of their command centers, artillery, and support units. Once this destructive mission had been completed, NATO's ground forces would return to their enclaves for reconstitution. Meanwhile, NATO's operational fires would begin attacking the enemy's second echelon, thereby starting the combat cycle over again.

This concept describes nonlinear NATO combat operations within a single corps. A similar pattern would be followed in other corps sectors under heavy enemy assault. Meanwhile, forces in corps sectors not under major attack would redeploy rapidly to more threatened areas or conceivably would launch supportive attacks into enemy territory. NATO's theater reserves would be employed in a similarly aggressive manner. Eventually the enemy presumably would be too battered to continue his attack. At this juncture, NATO's forces might swing over to conduct a theater counteroffensive in an effort to bring the war to a successful conclusion,

NATO's ultimate military goals, of course, would depend upon its larger political objectives, but there is nothing about this doctrine or the forces required to execute it that would prevent NATO from limiting itself to restoring the status quo ante on terms allowing for the safe withdrawal of Soviet forces. Indeed, this doctrine, in wholeheartedly embracing mobile warfare, potentially gives NATO's commanders a retaliatory and fatal knockout punch of their own. Without question, this doctrine fits comfortably within the framework of a purely defensive NATO military strategy. Nonetheless, it is well suited for being used offensively not only to restore lost territory but also to seek a classical military victory by destroying the enemy's forces. This is a capability that NATO's linear doctrine, whatever its other merits and demerits, never provided or sought to offer.

In moving NATO's doctrine away from positional warfare and toward mobile counterattacks, nonlinear defense has clearcut implications for NATO's force structure. It calls for NATO's ground combat forces to be configured as heavily armed units with high mobility, maneuverability, and firepower. It places a premium on armored and mechanized divisions equipped with large numbers of tanks, infantry fighting vehicles, and self-propelled artillery. It uses mechanized infantry within these units but seemingly provides little role for divisions that rely on infantry, truck transport, and towed artillery. An exception to this rule might be the idea of using a few light infantry units on favorable terrain, but this bow to infantry aside, nonlinear defense would perpetuate the armor-oriented philosophy that has guided NATO's ground posture for the past two decades.

Nonlinear defense also calls for continuity with respect to NATO's tactical air posture. It requires a powerful air force capable of conducting strike operations and providing direct support to maneuvering ground units. This concept would retain today's mix of interceptors, multimission fighter bombers, and specialized ground attack aircraft. For both ground and air forces, nonlinear defense would demand continued acquisition of high technology systems designed to upgrade NATO's firepower, survivability, and mobility. This would apply not only to new weapons and delivery systems but also to a broad array of C3I technologies, munitions, and support packages. Its adoption thus would have a stimulating effect on NATO's modernization programs and other improvement efforts.

An evaluation of nonlinear defense appropriately should begin with the larger policy and strategy issues that this doctrine raises. It would allow the Alliance's future military strategy to pursue the goal of a stalwart conventional defense even if NATO's forces are cut appreciably, by focusing NATO's employment concept on the task of defeating the enemy threat rather than protecting terrain. This focus liberates NATO from trying to maintain a

forward linear array with forces insufficiently large for their defensive mission. By moving NATO's center of gravity backward and relying on mobile operations to defeat the enemy, nonlinear defense reduces the heightened risks that otherwise would accompany force reductions.

Other things being equal, this doctrine probably would also enhance deterrence. A linear defense can wage a war of attrition but not annihilation. It might be capable of frustrating the enemy's intent, but it poses no active threat to his survival if he ceases his own offensive. A defensive doctrine anchored on mobile warfare can quickly cross over to the offensive. By giving the defender a knockout punch of his own, it can be employed actively not only to shatter the attacker's will to continue fighting but also to erase his forces from the battlefield. In creating the prospect that aggression might culminate in the physical destruction of Soviet forces, nonlinear defense seemingly adds an extra margin of uncertainty and cost to the deterrence equation. This feature would both strengthen NATO's overall military strategy and help compensate for whatever losses in psychological leverage are suffered by its disavowal of early nuclear escalation.

A drawback of nonlinear defense is that, by shifting NATO's defenses further to the rear, it might defend the Alliance's border regions less automatically and reliably than a strong frontal array. Since its goal is to prevent enemy forces from conquering and holding onto NATO territory, it is not inconsistent with the spirit and intent of the forward defense concept. But, like the doctrine of trading space for time, it does conduct forward defense more flexibly and contingently than the old doctrine. In particular, it might allow enemy forces to enter into friendly territory to set them up for NATO's counterstrokes. This scheme of operations leaves forward defense both dependent on NATO's ability to execute these maneuvers and politically vulnerable to a limited land-grab attack. In this sense, this doctrine trades off forward defense in favor of the military objectives of NATO's strategy.

Nonlinear defense probably would have a mixed effect on NATO's evolving doctrine for employing tactical nuclear weapons. By promising to make NATO's conventional posture more robust, it would enhance NATO's capacity to retire its nuclear weapons to a place of last resort. At the same time, maneuver wars are notoriously uncertain affairs that can be lost as easily as they are won. This doctrine therefore would offer no guarantees that nuclear weapons would never be needed. Moreover, this doctrine promises to make any employment of battlefield nuclear weapons more complicated than before. In a swirling maneuver battle, opposing forces might well be too closely interlocked to permit convenient nuclear targeting. This situation could increase pressures for initially crossing the nuclear threshold with deeper, more provocative strikes than formerly was the case.

Additionally, this doctrine might blur NATO's political criteria for escalating. It would make imminent conventional defeat less easy to measure, because a maneuver-capable defender normally still has options even if he is backed into a corner. Also, it might allow conventional combat operations to expand more deeply into NATO's territory than previously was contemplated. Both developments could complicate the already difficult task of deciding exactly when the nuclear threshold must be crossed.

Another troublesome issue is whether NATO should replace a doctrine that appears defensive with one than can be interpreted otherwise. For all its drawbacks, NATO's linear array clearly is tied to the goal of protecting Germany's borders and therefore is consistent with the Alliance's defensive strategy. NATO's adoption of a new doctrine that had the earmarks of waging mobile warfare with modern weapons would raise political eyebrows across Europe, including in Moscow. Its efforts to present a defensive image notwithstanding, even

NATO's traditional doctrine often was criticized by the Soviet government as being a wolf in sheep's clothing. Soviet complaints became particularly vociferous when AirLand Battle and FOFA entered NATO's thinking. The public unveiling of a new doctrine that seemed to carry NATO even further toward offensive concepts plausibly might not elicit a negative response if relations with the USSR continue warming. But if events take a different course as German unification and other developments take hold, the Soviets might choose to make an issue of the matter.

In any event, legitimate questions can be raised about whether a doctrine with an offensive twist squares with a NATO arms control policy that seeks to promote stability by reducing the capacity for offensive war. To be sure, NATO's overall strategy would remain defensive and its new doctrine would be consistent with CFE's force limits, yet there would be at least the surface appearance of inconsistency. Moreover, this doctrine could hardly be expected to encourage the Soviets to continue moving down the path toward a defensive doctrine that they only recently embarked upon. It might even help produce a future European force balance in which both sides are primed to wage mobile offensive warfare. In European history, military standoffs of this nature have not generally enhanced political stability.

The other side of the argument is that this doctrine clearly has attractive military features that should, if at all possible, be incorporated into NATO's plans. This doctrine is not an all-or-nothing proposition. In all likelihood, major parts of it can be adopted without transforming NATO's forces and strategy into an offensive fighting machine. In the end, the issue is likely to boil down to one of establishing an appropriate balance among competing priorities. When the Alliance was confronted with similar doctrinal innovations in the past, it normally succeeded in finding a proper blending of the new and the old. Typically it did so by adopting the new doctrine's most valuable elements while sanding off the rough edges. MC 14/3 is a good example, as is FOFA. If this past is prologue, nonlinear defense is likely to encounter a similar experience.

NATO could hardly hardly be criticized for embracing this doctrine if, as seems likely, the Soviets themselves preserve a healthy capacity for waging mobile warfare and conducting offensive campaigns. In this event, NATO's adoption of this doctrine might actually have a stabilizing effect by making its posture less susceptible to being defeated in the one kind of war still open to the Soviets. Additionally, the impending extension of NATO's defense umbrella eastward might well give rise to a need for forces that can fight on the move in meeting engagements. Finally, adoption of this doctrine might be made necessary by Alliance budgetary cutbacks or by the arms control process itself. All the other doctrinal options examined here, which preserve the framework of a linear defense, require some minimum levels of forces so that they can be confidently executed. Very deep cuts in NATO's posture, well beyond CFE I, could leave NATO's commanders little other choice than to mass the remaining forces together and to use them in mobile ways. At that juncture, NATO would have to embrace this doctrine to its fullest, regardless of any negative byproducts.

The Alliance is unlikely to be compelled in the immediate future to come to a firm conclusion about these tradeoffs. For the few years remaining in which NATO's defense plans will remain focused on the old inter-German border, member nations are not expected to reduce their forces to the point where adoption of this doctrine, rather than one of the others, is necessary. Future reductions on the order of 25 percent or more might change this calculus. But by that time, NATO's doctrinal debates probably will have shifted to the very different military and political issues raised by the task of protecting all of East Germany. The issue of nonlinear defense can then be settled in that context.

In the interim, NATO will have an opportunity to explore in depth the assets and drawbacks of nonlinear defense. One important issue will be whether all of NATO's separate national armies are capable of implementing it. Whatever its other qualities, nonlinear defense on a fluid battlefield is more difficult to execute than NATO's traditional doctrine. It requires considerable skill by senior commanders in orchestrating the constant motion of multiple combat units and support forces. It also requires modern weapons, including highly capable tanks, long-range missiles, advanced intelligence systems, and other technologies. Finally, it demands well-trained soldiers and a proclivity for flexible, independent actions by subordinate command echelons.

This doctrine seems ideally suited to the U.S. Army, which possesses the weapons, personnel, training regimens, and support forces required to perform maneuver operations. Also the U.S. Army has other theaters and contingencies to consider, many of which place a premium on mobile warfare and long-range firepower. Despite some deficiencies, the German and British armies also seem structurally capable of moving in this direction as they modernize, and French doctrine has long been oriented toward mobile warfare. Whether Dutch, Belgian, and Danish forces are similarly prepared is another matter. Since any layer cake array requires a common doctrine among its participants, the future of nonlinear defense might well be determined by whether these nations continue performing frontal roles in NATO's strategy.

A second concern is that a nonlinear defense doctrine is not likely to be purchased cheaply. It requires a suite of the most modern weapons available, all of which are expensive. Some of these systems have not yet passed through the various stages of research and development. This doctrine also demands the high readiness levels and training standards that are best provided by active forces, not reserves. Its logistic support and sustainability needs are not yet fully defined, but they are unlikely to be modest. Whether all this squares with an Alliance headed toward reserve forces, lower budgets, and modest acquisition programs is a noteworthy concern.

Another potential impediment is that NATO appears to be moving in the direction of establishing multinational formations. Evidently corps-sized units are to be created by drawing on division-sized national units. The primary motivation is the growing desire to establish greater political integration among NATO's forces. Unfortunately, this step is likely to have a negative effect on NATO's military preparedness, especially on its capacity to adopt a more difficult employment doctrine. Much will depend upon the degree to which the multinational concept is pursued. Modest steps, such as the formation of only a single NATO multinational corps, probably would have little bearing on NATO's doctrinal choices. But major steps, involving several corps, might drive NATO in the direction of preserving a simpler, more easily executed doctrine.

When the dust has settled, a final and perhaps determining issue will be whether a nonlinear defense in fact performs better militarily than alternatives that commit a larger number of NATO's forces to frontal positions. The drawbacks of a linear defense are well known. Over the years, many NATO officers have expressed their sentiments in favor of a mobile defense if only the political barriers against it could be knocked down. Mobile warfare, however, has drawbacks of its own that are commonly overlooked, especially for a multinational coalition made up of a large number of small armies with different weapons and tactics. Also, the specific weaknesses of the new nonlinear doctrine have yet to be unearthed. As a result, this doctrine benefits by what amounts to an unfair comparison.

The appearance of the nonlinear doctrine culminates a logical progression that began with AirLand Battle's rejection of concepts that ignored the operational level of war. It has come on the stage at a time when NATO's military debates had gone stale and the Alliance is casting about for a new military raison d'être. By adding a fresh perspective, it is prodding NATO in the direction of reappraisal and innovation. Most important, it has reintroduced military art into a calculus that had become too dominated by technology and mathematics. It has called attention to the fact that wars are determined by more than numbers and that military skill is an especially important variable. It has pushed NATO's defense plans from passivity into a reawakened awareness of the need to control a war by seizing the initiative. Also, it has helped resurrect the important role that offensive tactics can play in any well-construed defensive strategy. These are valid reasons why this doctrine has caught on so quickly in many U.S. and Allied military quarters.

Yet in military affairs, every action normally has an offsetting reaction and each doctrine carries with it the seeds of its own destruction. Once the nonlinear doctrine's problems are identified, the question will start being asked whether these problems are sufficiently worrisome to tilt the balance against it. The answer to this question currently is unclear, but this doctrine should be subjected to the same scrutiny that characterized U.S. and NATO debates over earlier doctrines that themselves offered high promise.

What are likely to be nonlinear defense's problems in the NATO setting? If a linear defense is weakened by its propensity to commit too many forces forward, a nonlinear defense—as defined here—might be undermined by its failure to establish a sufficiently strong frontal wall. This doctrine places faith in the ability of reconnaissance units, intelligence platforms, long-range missiles, and attack helicopters to perform many of the functions of a traditional front line. It calculates that firepower can substitute for mass and maneuver in many areas. It also assumes that the gains accrued from conducting flanking attacks will outweigh the losses suffered by no longer having prepared positions. Whether these propositions will turn out to be valid, especially once enemy countermeasures become known, remains to be seen.

Unlike its alternatives, a nonlinear doctrine offers the defender a potential military victory if he employs his counterattack options to best advantage. But this doctrine is risky and harder to execute and by seeking to strike hard against the enemy's flanks and rear areas, a counterattacking defender can expose himself to the same blows by the enemy. In a sense, mobile warfare is similar to a boxing match. Throwing punches is necessary to defeat the opponent, but the act of doing so can expose oneself to being knocked out in return. The art of mobile warfare, like boxing, requires dexterity and agility. These qualities are measured by more than simply throwing military punches; they are a product of a successful capacity to both thrust and parry, and they require maintaining an adequate defensive screen that can ward off the attacker's blows even as the defender launches counterthrusts of his own.[8]

The central idea behind maintaining a strong frontal wall in mobile warfare is to deprive the attacker of any opportunity to land a deadly blow. With the battle thereby brought under control, the defender can design his own offensive with confidence that attacking will not fatally expose himself. A purely linear defense takes this idea too far by committing so many forces to the frontal wall that insufficient units are left over to launch counterstrikes. A purely nonlinear doctrine runs the opposite risk of so denuding the frontal wall that the defender cannot parry the attacker's thrusts. In that case, the act of counter-

[8]For an insightful analysis of the problems associated with mobile defense, see Mearsheimer, 1983.

attacking could expose the defender in potentially fatal ways. The outcome of combat would be determined by who swings hardest, best, and last.

In assessing how these dynamics affect NATO's future doctrine, a critical issue is whether any future Central European war will be a parity battle or, as seems more likely, a situation in which NATO is still outnumbered by some margin. In the absence of a strong frontal wall, mobile warfare is best conducted when the defender and attacker are equally large. It is a dangerous doctrine when the attacker has larger forces because the attacker would still be capable of establishing a front wall of his own. In this situation, the attacker could win and the defender could lose, but not vice versa, which would be the case if the Soviets were to develop the deep fire and reconnaissance systems to offset NATO's advantages in these areas that the nonlinear doctrine heavily depends upon. Without any major qualitative edge elsewhere, NATO conceivably could achieve a stalemate by constantly maneuvering to block the Soviets' thrusts. But efforts to launch counterthrusts of NATO's own could be fruitless at best and disastrous at worst.

As long as NATO confronts the problem of fighting outnumbered by some margin, it will need to strike an appropriate balance between frontal power and rear area strength among its forward corps. With the modern weapons becoming available, a rearward-anchored nonlinear defense with few forces forward (25 percent or less) probably will be preferable to a linear array that places most forces forward (75 percent or more) and few back. But a different matter is whether a nonlinear defense will perform better than a doctrine that establishes an equal balance (50 percent apiece) between the two missions. A balanced array, for example, could establish the combination of forward strongpoints and counterattack forces that seemed to work well for the Germans on the Russian front. Complicated tradeoffs are involved here, and possibly the situation has changed since World War II in ways that alter this calculus, but military planners will need to address this issue as the nonlinear doctrine evolves in the years ahead.

Defensive (Nonoffensive) Defense

Defensive defense is the polar opposite of nonlinear defense. These two doctrines are harmonious only in the sense that both seek to compensate for the reductions in NATO's armor and other heavy weapons that CFE negotiations and Alliance budgetary cuts are likely to bring. Nonlinear defense would abandon NATO's frontal wall, defensive defense would seek to preserve it. Nonlinear defense would call on NATO's forces to prepare for mobile warfare, defensive defense would try to strip them of precisely that capability. Nonlinear defense would deploy armored and mechanized forces that provide high battlefield mobility, defensive defense would primarily deploy infantry forces that have little mobility. Rarely have two doctrines competing for NATO's attention been so completely different. Their divergent philosophies provide the Alliance a broad spectrum of choices in deciding how to prepare its future conventional defenses.[9]

The principal reason these doctrines differ so noticeably is that they are pursuing quite dissimilar strategic objectives. Nonlinear defense is preoccupied with the goal of preserving an adequate conventional posture as NATO's forces are reduced. It subordinates other goals to this task, and it discounts any risk that NATO's adoption of a mobile doctrine might appear to be politically provocative to other European nations. Defensive defense views

[9]von Bulow, 1988. See also articles in *Bulletin of the Atomic Scientists*, "A New European Defense," September 1988.

European security affairs through different lenses. Hostile to nuclear weapons, it believes that NATO should retain adequate conventional strength in the years ahead, but it also believes that NATO's own force posture and doctrine have played a contributing role in fostering East-West tensions over the years. Accordingly, this doctrine aspires to use force reductions and reform to pull NATO's allegedly offensive fangs, thereby promoting the cause of European stability. The package of doctrinal and force structure innovations that it offers is a product of its efforts to balance these two objectives.

Defensive defense aspires to field a conventional posture that can defend but cannot swing over to the attack, even in a counteroffensive. Under of this concept, various plans have been offered that would fundamentally recast NATO's posture and employment doctrine. There are, for example, the Afheldt, von Unterseher, and von Bulow plans, all named after the West German analysts who wrote them. These plans differ somewhat, and space does not permit a detailed evaluation of each of them. Instead, this document addresses their common themes, arguments, and recommendations.

In general, defensive defense is attracted to the idea of retaining a NATO linear positional array because this doctrine both protects German territory and appears less provocative than its more mobility-oriented counterparts. But defensive defense also acknowledges that deep force cuts could undermine NATO's conventional strategy by leaving the remaining posture too spread out. It proposes to solve this problem by retaining the Alliance's military manpower as tanks, infantry fighting vehicles, self-propelled artillery, strike aircraft, and other offensive weapons are removed. The resulting large manpower pool would be reconfigured as light infantry forces that could deploy into the forward areas and thereby recreate a strong frontal wall. These forces, however, would be so shorn of their tanks and other heavy weapons that they could not readily embark on a mobile offensive.

Defensive defense by no means would leave NATO's posture militarily impotent or completely denuded of armor. This doctrine essentially aspires to create a strong forward shield of infantry units backed up by a reserve armored sword of at least modest proportions. A more colorful label is that of a "spiderweb," a forward web of infantry positions backed up by small armored spiders capable of crawling over the web to reinforce threatened sectors. In theory, NATO's posture would be left with the very combination of a strong front line and reserves commonly called for by military analysts. This posture, however, would be armed very differently from the traditional model in ways that, by design, would render it largely immobile.

The forward wall would be populated by the equivalent of 35–45 infantry divisions (or more) organized into eight separate corps sectors. Some of these units would be active-duty formations but most would be mobilizable reserves. The composition of these units would be shaped by the goal of providing them with enough firepower and survivability to stop armor-heavy attacks. As a result, these units would retain the rifles, machine guns, mortars, towed artillery, and other weapons that commonly are associated with traditional infantry formations. They also would be equipped with a large number of modern antitank missiles, antiarmor drones, and possibly many short-range, single-warhead rockets for launching barrage fires.

Each division would be provided a small armor force of a single tank battalion (35–50) tanks, but it would have few infantry fighting vehicles. Its eight infantry battalions therefore would be straight-leg infantry units or, at best, motorized with trucks. This lack of mechanized infantry would dramatically reduce the tactical mobility of all the forward divisions. Equally important, it would prevent them from conducting offensive combined arms

operations. Their tanks could attack, but they would not be able to draw on the critically important support of mobile infantry and self-propelled artillery.

In compensation for this loss of mobility, the combat effectiveness of these divisions would be enhanced by prepared positions and measures to sculpture the terrain to help the defense. Prepared positions, possibly including fortifications and bunkers, would be constructed throughout the forward zone to a depth of 50–100 kilometers. This depth would confront the Soviets with the difficult task of slowly fighting their way through a thick wall rather than quickly punching through a thinner, more brittle line. The terrain-enhancement measures also would be ambitious. For example, a network of pipelines would be implanted that could be filled with liquid explosives in a crisis and then detonated to create deep antitank ditches. River banks and road shoulders would be cut at sharp angles to prevent enemy tanks from advancing over them quickly. Such obstacles as dragons'-teeth and modern mines would be purchased in large quantities. The goal would be to create a thick forward zone that enemy armor could not easily transit and NATO infantry units could readily defend.

Backing up this forward zone would be the equivalent of about ten divisions of armored and mechanized units. These forces would be armed along traditional lines, with about 75–100 tanks per brigade and sufficient infantry fighting vehicles and self-propelled artillery to tubes to create a combined arms team. Their mission would be to provide a mobile reserve for the forward units. These units would neither be massed together nor deployed far back in the rear areas in the manner that would enable them to launch large counterattacks across the entire theater. Rather, they would be distributed evenly along the front in small doses and assigned directly to the commanders of the forward corps. They thus could launch small attacks locally, but they could not conduct large, deeply penetrating counteroffensives.

The act of transforming NATO's posture in these ways would dramatically reduce not only its offensive capacity but also its overall combat power. NATO's posture at M+30 presumably would still field about 45 DEs as measured in maneuver units and combat manpower, but its weapons would amount to only about 20–25 DEFs of strength, or about 50 percent of the current total, well below CFE-authorized levels. In particular, it would have far fewer heavy weapons. The present posture fields about 10,000 tanks, a similar number of infantry fighting vehicles, and about 4000 artillery tubes. The defensive defense posture would deploy only about 3000–5000 tanks, even fewer infantry fighting vehicles, and 2000 artillery tubes. This posture presumably would have enough mass, firepower, and survivability to defend the Central Region terrain in adequate density and strength, but it would lack the heavy weapons, mobility, and combined arms capability to execute any doctrine other than a stationary defense.

Defensive defense would pull NATO's offensive teeth in other ways. The headquarters staff of each corps and division would be kept small and would lack the manpower and communications assets to conduct coordinated maneuvers with its subordinate combat units. The frontally committed divisions would be deployed in considerable depth rather than bunched together. This pattern would prevent their brigades and regiments from quickly joining together to provide each other the mutual support that is needed in attacking. The prepared positions of these units would be linked by underground wires, thereby preventing them from communicating with each other if they abandoned their positions to attack the enemy. Additionally, logistic support forces for both the forward infantry divisions and rear armored units would be kept small and denuded of the engineer, maintenance, and ammunition hauling units that enable a military force to move forward.

These constraints are the intentional product of a conscious design concept. Taken together, they would go a considerable distance toward locking NATO into a stationary theater employment doctrine. Their effects initially would be physical, but later they would expand into NATO's military psychology. NATO's tactical doctrine, training standards, exercise regimens, and even weapons procurement eventually would be drawn into conformity with defensive defense's dictates. In the end, NATO's forces would be hard-pressed to move forward or otherwise engage in offensive adventures even if the Alliance's political leaders gave them the order to do so.

Defensive defense entered the public limelight in 1987–1988 when arms control negotiations were the primary hope for ending the NATO-Warsaw Pact standoff at the inter-German border. It has faded from public debate since then because larger political events are accomplishing this goal, but it is likely to reemerge as one alternative for bringing stability to the European security system that looms ahead. Consequently, it still merits careful scrutiny, but with one eye on the old inter-German border and the other on the military missions ahead.

Defensive defense has been criticized vocally by several NATO senior military officers, including General Galvin. Although many charges can be levied against it, its merits deserve acknowledgment. Perhaps its chief strength is that it does ask the right question: How can NATO confidently defend itself, protect its exposed border regions, and behave nonprovocatively all at the same time? Its capacity to focus on the correct issue is one reason for approaching this doctrine in a positive spirit that attempts to sort out the good from the bad.

Defensive defense helps establish the important point that the military balance in Central Europe will become more stable if all participants exert a stronger effort to design their forces for defense, not offense. It also recognizes that while a NATO frontal wall is an important feature of a defense strategy that needs to protect exposed borders, a thin array can be beaten even at parity. It is not blind to the problems that can arise if NATO's posture, in overall size, is cut too far in the absence of a major change in force structure and employment doctrine. Moroever, it usefully alerts NATO to the risk that any solution relying too heavily on mobile warfare might be misconstrued by nations that are an important part of the stability equation. Further, it does provide NATO one path for pursuing deeper cuts in arms control negotiations and for adjusting to unilateral reductions that will be coming.

Militarily, defensive defense offers many helpful ideas on better using infantry forces, reserve mobilization, low technology, prepared positions, and terrain development. All of these areas have been chronically overlooked by a NATO defense philosophy that has emphasized active units, mechanized formations, high technology, and open terrain. Even if defensive defense is not adopted as a whole, several of its specific measures probably will be appropriate for NATO's future strategy and therefore should receive careful attention.

On the negative side, defensive defense suffers from a critical drawback that grows directly out of its core theory. In its quest for stability, it carries too far the idea of paring back NATO weapons and force structures that it regards as offensive. In the process, it strips away systems that are vitally needed for defense and especially for executing the forward defense that it endorses. To compound matters, it overestimates the combat capability of the infantry units that it would leave behind. The force posture and employment concept that it endorses would be nonprovocative, but they would not pass a more important test of sufficiency, the confident ability to defend successfully. In failing to meet this test, they would undermine this doctrine's goal: military stability by ensuring that neither side has the capacity to overwhelm the other.

Defensive defense too readily accepts the assumption that NATO's present posture has the trappings of an offensive strategy. NATO's forces, it is true, are well-armed with weapons that can be used in offensive missions, but weapons are only one element in the military equation. With only 45 divisions at D-Day, NATO's posture falls far short of the standards that commonly have been associated with a theater offensive campaign. Over the years, it has been the Soviets, with 80–100 mobilizable divisions, who have met these standards, not NATO. Moroever, NATO's logistic support structure already is itself sized for defense, not forward-moving offensive operations. This especially is the case for the German Army, whose support forces from inception purposely were kept small to signal defensive intent. Consequently, defensive defense misconstrues the Central European military balance. This erroneous view leads it to endorse major reductions of weapons whose main effect has been to help make a defensively sized NATO posture capable of performing its mission.

The force posture that defensive defense advocates might be adequate for its task if the adversary could be relied upon to mimic it. In that case, a stable situation could arise because two immobile infantry armies with only small armored reserves would be confronting each other on densely covered terrain. Precisely because neither side had sought to develop an offensive capability to use against the other, defense would be the only military option open to each. This doctrine thus fits neatly into a European security system that already is at peace, with no prospect of war between the Soviet Union and NATO.

This sort of future political consensus in Europe, however, is not something that can be taken for granted. The proper question to be asked is whether this doctrine fits into a European security system that still suffers from sufficiently serious tensions to keep the possibility of aggression and war alive. In particular, what kind of situation would arise if NATO were to adopt this doctrine unilaterally while the Soviets remained inclined, subject to CFE's constraints, to preserve a healthy set of offensive capabilities for themselves? NATO could find itself confronting a still mobile Soviet military force that is well-endowed for ripping an immobile infantry defense to shreds. To compound matters, NATO would be trying to defend with a posture and employment scheme strongly resembling what the French Army erected against German aggression in the spring of 1940.

Present trends suggest that although the Soviets are shifting their posture in the direction of a defensive orientation, they are not denuding themselves of offensive options. Some of their divisions evidently are being configured as artillery-rifle (infantry) units, but the bulk are to remain as tank or mechanized divisions. The tank inventories of these units are being reduced, but both types of divisions will be left with plenty of armored and mechanized weapons to conduct mobile warfare. Moreover, they will be capable of fielding enough divisions to attain at least some margin of numerical superiority over NATO, thereby further tilting the balance in their favor.

These developments bode ill for defensive defense. After adopting this doctrine, NATO would face a situation of near parity in combat manpower but a numerical imbalance of about 1.2–1.6:1 in maneuver units. The weapon reductions mandated by defensive defense would leave NATO outnumbered in total firepower (DEF scores) by 2–3:1. This is far too large an imbalance, especially since NATO would almost completely lack mobility, could not conduct a combined arms campaign of its own, and could not counter-concentrate in anything like a timely fashion.

Under these conditions, NATO's frontal array could be hard-pressed to contain a concentrated Soviet attack. Its frontline infantry units would have enough antitank weapons to destroy Soviet tanks at long ranges, and their prepared positions would give some protection

from returning fires; but they still could be suppressed by enemy artillery and infantry, and their lack of tanks would prevent them from generating the high antiarmor fire rates that are critical as Soviet tanks draw near. Unable to win the close battle because their antitank missiles could not fire fast enough, they would be compelled to retreat. The absence of infantry fighting vehicles, however, would prevent them from leaving their positions safely. The likely result would be not a stalwart defense but rather their quick destruction.

As the battle passed through the frontal zone and toward the rear, NATO's fortunes probably would take a turn for the worse. Since defensive defense would have left NATO's infantry positions deployed in depth and therefore spread apart, they would be vulnerable to being picked off one by one. Their weak command staffs would inhibit their efforts to work cooperatively, and their lack of infantry fighting vehicles would prevent them even from massing together. The absence of enough tanks, in turn, would prevent them from conducting sufficient tactical counterattacks to disrupt the enemy's tempo and slow his momentum. This dynamic could produce a cascading defeat in detail that NATO would lack the physical capacity to arrest even if it abandoned defensive defense and tried to improvise.

NATO's ability to prevent an enemy armored breakthrough that almost inevitably would be fatal to an infantry defense would hinge on its ability to counter-concentrate forces in the main sector. The Soviets would not be able to punch through NATO's thick frontal zone immediately, but time here would be measured in relative terms. Because NATO would lack massed armored reserves in the rear areas, its own response probably would be sluggish, uncoordinated, and haphazard. And even if NATO responded promptly, it would not be able to assemble large enough armored and mechanized units to balance the kind of breakthrough force that the Soviets could marshall. Unable to draw on its infantry units committed to fixed positions elsewhere, NATO could find itself fighting a swirling maneuver battle in the rear areas and locally outnumbered by 5–10:1. This scenario easily could lead to the same kind of disaster that befell the French in 1940.

This grim portrayal does not imply that the doctrine of defensive defense is inevitably condemned to failure, but it does mean that this doctrine is unlikely to provide a confident defense with the level of NATO forces that realistically will be available. Defensive defense could work if its forward zone was sufficiently thick at all points to wear down a concentrated enemy attack before his forces could gain a breakthrough. A zone this thick, however, could be established only with an infantry-dominated force posture that is considerably larger than what NATO can now field. Precisely for this reason, some defensive defense schemes have called for an actual buildup of NATO's forces at D-Day to a level of 55–66 divisions, well beyond present plans. Individually, these divisions plausibly could be purchased more cheaply than their armored counterparts, but the total budgetary cost would remain high and, in any event, manpower requirements would rise. Given these resource demands, NATO's willingness to answer this doctrine's call anytime soon will be, at best, questionable.

Defensive defense also would require a concerted effort to prepare the terrain along Germany's borders by building an extensive network of concrete positions and making other preparations. At the height of the Cold War, these measures proved unacceptable to the German government, and they are even less likely to be acceptable in the years ahead. Moreover, a Soviet attack would still progress a good distance into German territory before grinding to a halt. Defensive defense schemes commonly call for a forward zone 100–150 kilometers deep, well beyond the depth of NATO's present plans. It therefore fails to pass the test of political acceptability.

The theory of defensive defense looks attractive at first blush, but it breaks down when battlefield dynamics and offense-defense interactions are examined in detail. It erroneously regards mobility, combined arms weapons, centralized command structures, modern communications systems, and large support forces as being needed only for the purpose of waging offensive warfare. These capabilities are indeed required for going on the attack, but they also are critical for defending successfully on the modern battlefield, especially against a well-prepared opponent. As a result, the idea of resorting to immobile infantry to form NATO's frontal wall would compound the Alliance's defense problems, not solve them.

Another argument against tilting heavily in the direction of stationary infantry is that Germany's unification might well make mobility an even more important factor in NATO's defensive calculus. NATO's major combat formations are likely to remain based in western Germany but could have wartime assignments in eastern Germany. Further, they might have to move there quickly and conceivably fight meeting engagements once they arrive. Beyond question, this situation would require armored and mechanized forces. Immobile infantry units would be even less appropriate there than along the inter-German border.

Defensive defense thus wrongly concludes that major surgery on NATO's posture can safely be undertaken with the aim of removing only offensive components while leaving the defensive systems behind. In reality, offense and defense are too intertwined to permit wholesale changes of this magnitude. Minor surgery, however, might be possible. For example, a moderate reduction of NATO's tank forces is one possibility. A modern mobile division normally requires more than the one battalion of tanks that an infantry unit would field and that defensive defense envisions. The present structure of four or five tank battalions is not necessarily a sine qua non for units performing frontal roles. The emergence of lethal antitank missiles is reducing, to some degree, the need for the rapid gunfire and kinetic energy penetrators that the tank provides.

A logical innovation would be to alter NATO's frontally committed mechanized divisions by reducing their armor inventories to three tank battalions apiece. This change would produce a standard division of three mechanized infantry brigades rather than the current configuration of one armored and two mechanized brigades. Similarly, it might be possible to alter NATO's armored divisions by reducing their armor to four tank battalions. This would yield a new division of one armored brigade and two mechanized infantry brigades, rather than vice versa. The withdrawn tanks, of course, would have to be replaced by infantry fighting vehicles to mechanize the infantry battalions that would take their place. The idea would be to leave NATO's heavy divisions with enough armor and mechanized infantry for their defensive tasks while taking the sharp edge off their ability to conduct deep mobile operations.

This innovation would have to be studied carefully to ensure that NATO's divisions remained properly armed to perform the new missions that are likely to arise in Central Europe. One useful indicator is that the Soviets are reducing the tank inventories of their divisions to these levels: 150 tanks for a motorized rifle division and 260 tanks for a tank division. To the extent that the Soviets are setting the pace, this level of armor might be enough for the divisions of the future. If this change were implemented across NATO's entire posture, the overall result would be to reduce NATO's armor requirement by 20 percent, from 10,000 tanks to about 8000 in combat units. This level is still well above the 3000–5000 tank posture advocated by defensive defense. Nonetheless, it does amount to an appreciable downward shift in NATO's tank needs, one that would help blunt any perception that NATO's forces are too heavily tilted toward offensive warfare.

Similar steps might be possible in other areas. Introduction of some towed artillery tubes and wheeled armored vehicles are examples of how offensive power might be scaled back, defensive capacity preserved, and budgetary costs reduced. The acquisition of more attack helicopters is another development that will help permit reductions in armor and offensive mobility. In the long run, the emergence of new defensive technologies might make possible further cutbacks in armor. For example, the development of high energy guns might permit the design of lighter tanks with less of the frontal armor that is associated with offensive tactics. The acquisition of mobile, gun-mounted tank destroyers is another possibility. In assessing these alternatives, NATO's forces will need to retain a properly balanced combined arms capability. But as long as this standard is met, NATO might able to take useful steps in the directions urged upon it by defensive defense.

Toward a Synthesis

NATO has available a variety of alternatives for adjusting its employment doctrine and force posture to compensate for the kinds of force reductions that are likely to occur in the near future. It will be hard pressed to adopt any one of these alternatives in a complete and exclusive way because all of them bring problems and barriers in their wake. This constraint will compel NATO to develop a package of measures that draws on each of these alternatives in the limited ways that will be possible. Fortunately, a combination of initiatives should be adequate to shore up NATO's conventional defenses as long as member nations do not make major cuts in their force contributions.

Shifting NATO's posture northward provides the greatest leverage in maintaining an adequate defense in NORTHAG. This step makes even greater sense when it is viewed as a forerunner to a subsequent effort to refocus NATO's plans on defending eastern Germany after Soviet forces are withdrawn. Neither the United States nor France is likely to be willing to spend the funds required to move its bases and supply dumps northward, nor would Germany itself be tolerant of the disruption that this change would cause. But these nations can at least develop operational plans to move these forces in a crisis. This step alone would help identify and overcome some of the barriers that might arise in an actual crisis. A similar conclusion is applicable to the idea of improving NATO's capacity to counter-concentrate. Even where programs are not possible, plans can carry NATO part way to the objective. With respect to trading space for time, NATO cannot politically afford to contemplate any wholesale disavowal of its forward defense goal. Expediency nonetheless dictates that limited retrogrades should be incorporated into NATO's plans on a contingency basis. Figure 13 illustrates how these measures might be combined to bolster NORTHAG's defenses.

NATO's military planners clearly are moving toward the idea of a nonlinear defense. This step is a healthy response to the changes at work in NATO's posture and on the modern battlefield. The exact manner in which nonlinear defense comes to be defined, however, matters greatly. The tradeoff between frontal strength and rear mobility needs to be considered in the context of both the inter-German border and what is likely to become a new NATO defense line on the Oder-Neisse rivers. It also needs to be addressed in relation to the kind of force posture that NATO will want to deploy and will be able to afford in the years ahead. Nonlinear defense implies mobile warfare, heavy forces, high readiness standards, and costly military programs. For political reasons, however, NATO is committing itself to a more transparently defensive stance. Also, the European allies seem headed toward reserve forces, lower budgets, and perhaps a greater reliance on infantry. These trends clearly will pull NATO away from nonlinear defense.

Fig. 13—An Improved NORTHAG Defense Plan

A proper balance between these competing stances might emerge in the context of a NATO effort to develop an altered division of military roles and missions in Central Europe. This could involve a movement by the German Army toward defensive defense in some respects, accompanied by the evolution of the U.S. Army toward nonlinear defense. The difference between them would be a matter of degree, not type. Both armies would remain mobile and well armed; however, the former would be designed more for the tactical mobility that accords with frontal defense missions, and the latter would aim for the operational, theaterwide mobility that is needed for rear area missions.

Especially since NATO's employment doctrine will be shifting northward and eastward, the Germany Army will continue to perform the frontal defense mission. Political considerations will dictate that it preserve its purely defensive appearance, and internal budgetary pressures will lead it to lean more heavily on using infantry forces for some functions. At the same time, it will need to remain highly mobile and capable of maneuvering on the battlefield. A comprehensive switch to defensive defense would not make sense, but a shift in emphasis away from armor toward mechanized infantry and helicopters might be appropriate.

The U.S. Army, meanwhile, will probably revert more to acting as a backup mobile reserve, and it is less likely to be subject to political design criteria and budgetary constraints. These conditions will encourage a greater emphasis on armor, other heavy weapons, and the mobile warfare concepts that normally accompany them. A NATO defense concept that encourages the German Army to adopt some defensive defense measures while the U.S. Army configures itself for nonlinear defense might make both military and political sense. Whatever the case, this complex issue will need to be addressed in the framework of a NATO employment doctrine that will be moving eastward.

DEFENDING EASTERN GERMANY: THE TRANSITION PERIOD

Since Soviet forces will be departing shortly, "end game" issues will play by far the largest role in shaping NATO's long-range employment doctrine and force posture for defending eastern Germany. However, the transition period ahead poses important, if temporary, challenges of its own. It will give rise to new demands on NATO's defenses that are a forerunner of what the Alliance will face, in more enduring ways, after Soviet forces have entirely withdrawn. In particular, it draws attention to the problems that would arise in mobilizing and moving NATO forces eastward at all future stages.

An analysis of NATO defense issues in this period must first postulate the manner in which an intense political crisis might escalate to war. It is difficult to see how NATO and the Soviet Union could come to loggerheads in Central Europe anytime in the next few years. Nonetheless, it is at least plausible—for planning purposes—to imagine a crisis unfolding over a Soviet refusal to leave on schedule and escalating thereafter into a larger Soviet threat to Germany's sovereignty and territorial integrity. That scenario is used here.

A showdown of this sort would be wholly unlike anything that was experienced during the Cold War. Militarily it would be entirely different because NATO would heavily outnumber the Soviets, who would have only a small force remaining in Germany. In contrast to years past, the Soviets would lack the wherewithal to quickly begin large-scale offensive or defensive operations. In all likelihood they would have to retreat into a defensive enclave along the German-Polish border and wait for reinforcements to arrive from the USSR. Politically this scenario would be unique because NATO no longer would be defending land that its forces already occupied in strength. NATO would now face the more complex problem of trying to extend protection over exposed Alliance territory that already had enemy forces deployed on its soil. To put matters mildly, this situation would pose unique demands on NATO's forces, crisis-management policies, and mobilization procedures.

How should NATO's present plans be adjusted to provide an appropriate spectrum of options for dealing with this situation? Conceivably NATO would be so afflicted with caution and uncertainty about Soviet intentions that it would keep its forces in western Germany. Provocative behavior by Soviet forces in eastern Germany coupled with mobilization in the USSR, however, would be an unmistakable sign that a military crisis was in the offing. In this situation NATO might want to preempt an impending Soviet reinforced attack by taking proactive deployment measures of its own.

This step could be motivated by a political desire to signal resolve. Ideally NATO might want to induce Soviet forces to leave Germany entirely. Reluctance to trigger a war, however, would weigh against any NATO attack on them. More realistically, NATO might want to help assert Germany's sovereignty by establishing firm military control over as much eastern German territory as possible. Reinforcing this calculus could be a military desire to

establish better defense positions on more advantageous terrain than what exists along NATO's old defense lines in central Germany. Whatever the motives, the net effect could be to induce NATO to mobilize its forces and move them some distance into eastern Germany well before any fighting actually begins.

Employment Concepts

For geographical reasons, two employment concepts beckon as a logical focus for NATO's defense plans during this interim period. They are displayed on Fig. 14. The first, more modest concept envisions the establishing of a NATO defense line primarily on the Salle River, a distance roughly 100 kilometers into eastern Germany. The second concept is more ambitious. It calls for a NATO defense line running down the Elbe River, which bends some 200 kilometers into eastern Germany south of Berlin. Both are attractive politically because they would assert a NATO crisis presence over substantial portions of eastern German territory. They would make sense militarily because they would enable NATO's plans to

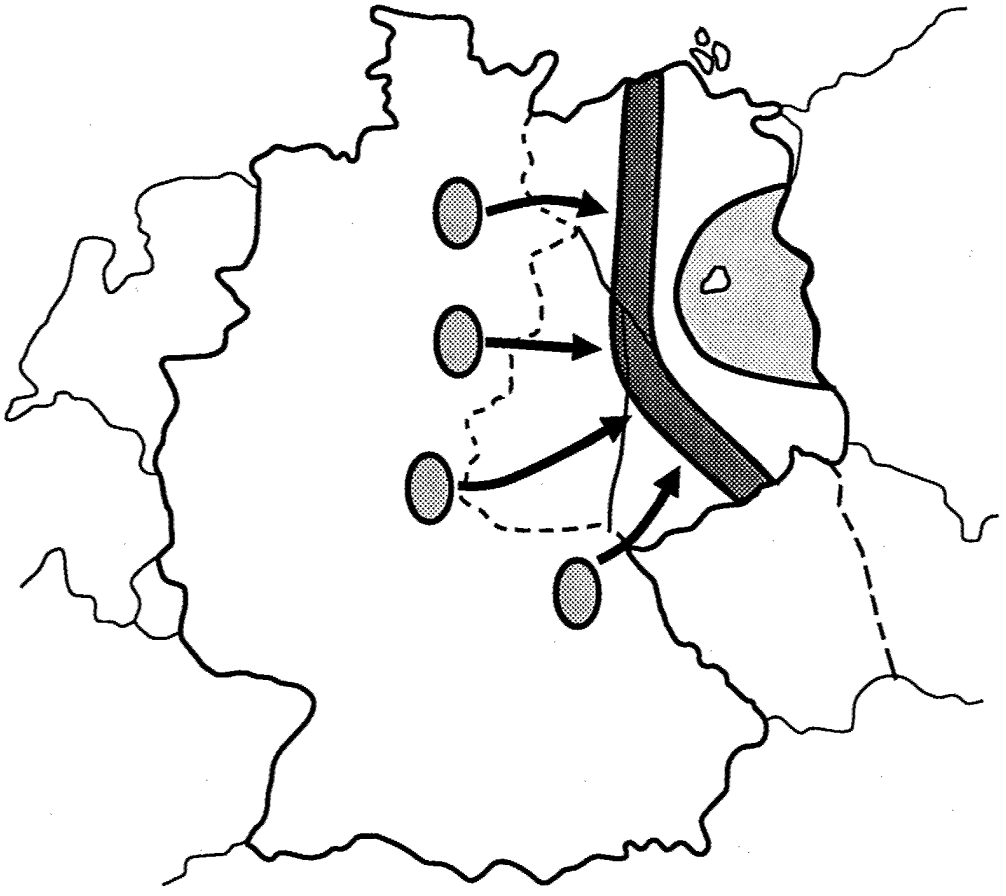

Fig. 14—Elbe Defense Line

concentrate on a 400 kilometer line that would be established behind one of two defense-favoring rivers.

Both concepts will be compatible with the NATO ground posture that will be available over the next few years, especially if NATO shifts its center of gravity northward. Of the two concepts, the Elbe line is preferable on both political and military grounds, but it would draw NATO's forces close to a Soviet enclave. Both concepts would enable NATO to defend more strongly than at the old inter-German border with its open terrain. NATO would be able to anchor its initial military plans on rivers that would be hard for even a reinforced Soviet force to cross quickly and easily. This would help strip away the Soviets' advantage in surprise and concentration and enable NATO to withhold more forces as reserves. Both concepts would probably not be implemented in a traditional linear fashion, but instead with some combination of the doctrines discussed earlier. Whatever doctrinal mix is chosen, NATO would be better able to achieve the blend of strong frontal defenses and mobile reserves that is the best guarantee of success.

Need for Planning

A NATO effort to establish either a Salle or Elbe defense line would be a complex undertaking that could become hostage to a set of impediments and constraints. These problems could be especially troublesome if NATO had to move quickly and under ambiguous circumstances. Some of them appear beyond remedial action, but others could be ameliorated by NATO planning aimed at smoothing out some of the rocks along the road.

Planning especially would be needed because NATO has no historical experience in this kind of operation. For the past 40 years, NATO's military strategy has been focused largely on protecting territory it already controls with forces that are located nearby. The 1949 Berlin airlift is one exception to this rule, and a decade later the Berlin crisis led NATO to examine military operations for liberating that city. That review ended with the conclusion that apart from symbolic gestures, NATO had few options at its disposal. Since then, NATO's defense plans have been mostly reactive. And apart from the problem of deploying NATO forces into Norway's northern regions, NATO has not had to deal with the task of moving large forces cross country over long distances. For this reason alone, eastern Germany's entrance into the Alliance necessitates an effort to rethink traditional ways of planning in Central Europe.

Political Requirements

A NATO effort to move eastward in a crisis easily could run afoul of the Alliance's own inability to make a clear political decision to take this step. Any crisis drawing NATO's forces into eastern Germany while Soviet forces were still based there would not unfold suddenly. It would grow out of a lengthy deterioration in political relations marked by NATO efforts to resolve the crisis diplomatically. Even so, the situation unavoidably would be highly ambiguous, and there would be major uncertainty about how Western actions might affect the Soviet Union's behavior. This easily could give rise to confusion and hesitancy on NATO's part.

This hesitancy could interact with NATO's brand of centralized, consensus decisionmaking. Particularly since an agreement has been reached not to deploy NATO forces into eastern Germany in peacetime, a formal alliance political decision would have to be made to take this step in a crisis. Under NATO alert procedures, SACEUR has the legal authority to

declare a state of Military Vigilance, thereby allowing him to take some military precautions. He has the authority to recall active duty troops to their bases, increase the alert status of air defense forces, and deploy limited ground forces (e.g., armored cavalry units) to the border areas. These steps would leave him well short of the capacity to begin moving large forces eastward. To put NATO's forces into a position to undertake major military operations of any sort, SACEUR would have to call a Simple Alert followed by a Reinforced Alert. These two alerts would enable him to initiate the numerous measures, including the transfer of authority over member-nation combat forces to him, that are required for this purpose. Both forms of alert, however, can be authorized only upon a political decision taken by the governments of NATO's nations.

Whether NATO's governments would be willing to act expeditiously would depend upon the situation at hand. NATO's alert procedures, and the institutional and political mechanisms supporting them, were established on the principles that NATO would be protecting territory it was already holding and that a Soviet attack was unambiguously about to be launched. Even so, NATO exercises over the years typically have seen the alert process take several days and weeks to unfold, with numerous delays caused by political hesitation.

Motivating NATO's governments to move into eastern Germany could be marked by even greater uncertainty, debate, and procrastination. In the interim, NATO's window for acting successfully might have begun closing if Soviet reinforcements were deploying westward from the USSR. Although NATO planning could not eliminate this risk, it might help reduce it. In particular, it could help sensitize NATO to this new issue, define a spectrum of responses, and establish criteria for determining what role force deployments should play in NATO's crisis-management policies. At a minimum, steps like these could clarify matters, save valuable time, and reduce NATO's need to improvise at the last moment.

Movement Dynamics

Assuming a NATO political decision is made to move eastward, NATO's forces would have to be capable of executing this maneuver in a timely and effective fashion. In all likelihood, NATO would have to move quickly after both sides had begun mobilizing, within the first week. At that juncture, NATO would have about 35 DEFs theoretically at its disposal. On the surface, this posture appears amply large for the mission, but more is involved than simply generating forces. Detailed decisions have to be made in securing their commitment and coordinating their movements. It is here that planning could have its most positive effect.

One of the most important issues is that of deciding what forces would be earmarked to carry out this maneuver. Would this movement best be undertaken by only German forces or should it be carried out by a NATO force composed of units from several nations? Arguments can be made in both directions. A purely German maneuver might be less provocative since the Kohl-Gorbachev agreement allows that nation to deploy its own forces into eastern Germany. This step, however, could leave Germany both out on a political limb and lacking adequate forces for the task. A multinational commitment would signal greater NATO unity and permit a more adequate posture but could provide the Soviets an undesirable pretext. NATO probably would be best advised to develop both types of options, but it would need to have addressed this issue in advance.

Quite apart from what is preferable, which nations actually would be willing and able to participate? Germany clearly would rank high in both attitude and capability. The U.S. and

British forces still deployed in western Germany—five or six divisions in 1993, but declining to about four divisions or less afterward—also are likely candidates. The stance of other NATO members might be less certain, especially in the early stages of a crisis. France traditionally has shown an aversion to acting as a member of any NATO integrated military operation. A host of political and military constraints could lead Belgium, the Netherlands, and Denmark to demur initially as well. If so, NATO could find itself with a committed force of only about 20 division equivalents, reinforced shortly thereafter by several U.S. divisions from CONUS. This posture probably would still be adequate for a Salle or Elbe defense (or along the Oder-Neisse), but it would be less robust than had all of NATO's forces been committed.

NATO's posture today is so highly integrated that failure by some member nations to participate might inhibit any effort to move into eastern Germany by the forces that would be available for this mission. In principle, German, U.S., and British units could extract themselves from their entanglements with other Allied forces and operate as a separate entity. In reality, this task could be complex and more difficult to accomplish. For example, senior Allied officers from nonparticipating nations hold positions at SHAPE, AFCENT, and other NATO command staffs. Would they join in planning for the move eastward even if their parent nations chose not to commit their own combat forces? Beyond this, soldiers from these nations operate NATO's radars, communication nets, command and control networks, and air defense sites. U.S. combat forces rely on support contributions from most of these nations. The naval forces of these nations also play important roles in NATO's maritime posture, especially in protecting convoys approaching European seaports. A NATO move eastward might be difficult without at least the passive acquiescence of these nations.

Another important issue is determining how NATO's ground forces would physically enter eastern Germany. The peacetime disposition of NATO's ground forces probably would play a major role in this decision, as would the layout of autobahns leading into eastern Germany. The interaction between bases and autobahns would push NATO toward entering eastern Germany primarily from the south. Military exigencies, however, might compel NATO to establish greater strength in the north by shifting forces there.

As Fig. 15 shows, NATO's forces currently are based in a linear fashion up and down the countryside in western Germany. Near Hamburg in the far north are a German division and a well-armed Territorial Army reserve brigade. Between Hamburg and Hannover, under NORTHAG's command, are German I Corps with four divisions and an airborne brigade and British I Corps with three armored divisions. To the rear is a U.S. brigade based at Garlstedt. Further south are CENTAG's forces. Near Kassel is German III Corps with three divisions, an airborne brigade, and two Territorial Army brigades. Between Fulda and the Czech border are the powerful U.S. V and VII Corps with four divisions, two separate brigades, and two armored cavalry regiments. In Bavaria is German II Corps, with three divisions, an airborne brigade, and two Territorial brigades. Far to the rear by the Rhine River is the French First Army. The figure does not show Belgian, Dutch, and Danish forces, nearly all of which are based on their home territory.

For the next two to three years at least, this basing pattern probably will shrink somewhat but remain fairly constant in its geographic distribution. There will be two to three fewer U.S. and British divisions deployed on German territory, and some German units will have been converted to reserve status, but the north-south layout of NATO's bases is unlikely to change a great deal.

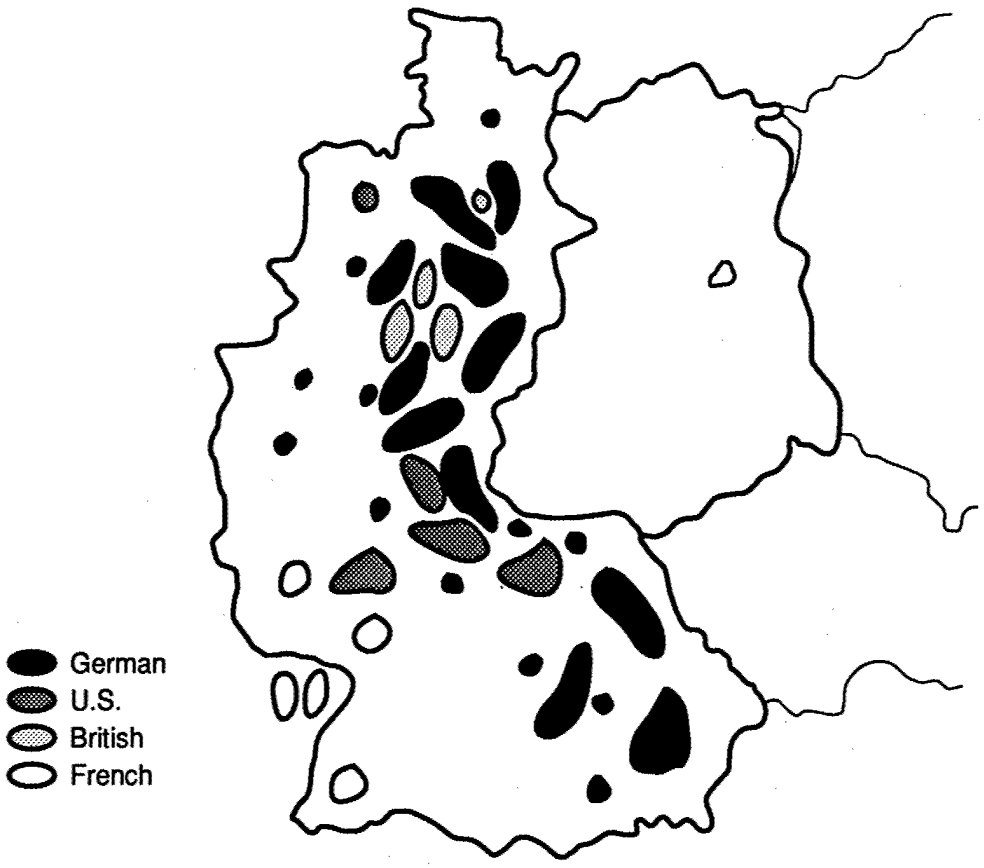

Fig. 15—Peacetime Basing Locations of NATO's Ground Forces

Assuming only German, U.S., and British forces are involved, how quickly could they begin moving eastward and in what formation would they deploy? Provided these forces are properly trained and manned, nearly all could be ready to move out of their bases within a day or two. The task of assembling them into a formation conducive to traffic control, however, would take about another three days. As a result, they probably could not begin crossing into eastern Germany in strength until 5–6 days after the order had been given. Their axes of advance would probably be determined by the location of their bases in relation to the four autobahns that lead to Berlin.

Figure 16 suggests that NATO's forces initially would be deployed with the weight of emphasis on the southern two autobahns. This formation would not be optimal for NATO because it is so heavily tilted to the south. On the northern flank, a fairly small force of eight divisions would advance to establish blocking positions on the 200 kilometers of open terrain between Hamburg-Hannover and Berlin. On the southern flank, meanwhile, ten divisions would advance eastward to guard an area that is less threatened. NATO could circumvent this maldistribution by transferring four divisions from southern Germany, as displayed on Fig. 17. Most likely another four or five days would be needed to shift not only the necessary

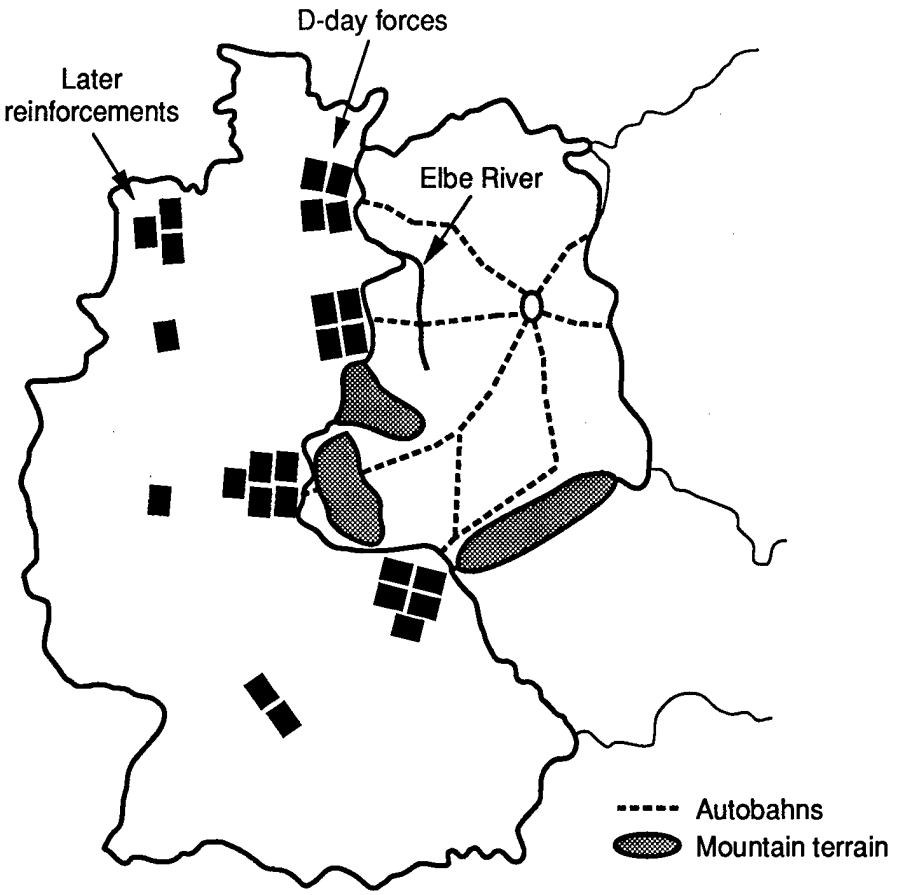

Fig. 16—A Southern-Oriented NATO Advance into Eastern Germany

combat forces but also their support units. It would need to be done if a properly strong NATO defense line were to be established on the Salle or Elbe rivers.

In addition to mobilizing its active forces for this mission, NATO also would have to mobilize sizable reserves, particularly for the German Army. Currently the Field Army's 12 combat divisions are manned primarily with active troops and require mobilization of only about 50,000 reservists to reach authorized manning levels. Many of the Field Army's combat support and combat service support units, however, rely heavily on reserve manpower to reach full strength. The Field Army's wartime authorization is 50 percent reservists, all of whom would have to be called to active duty in a crisis. In addition, Germany would mobilize its reserve Territorial Army of several hundred thousand soldiers, 85 percent of whom are reservists. This force plays an important role in Germany's defense plans, including providing support for the Field Army. Additionally, it provides about 100,000 troops that perform important host nation support functions for U.S. forces.

Logistic constraints also could complicate matters. For years NATO has judged that the support variable would work in its own favor. One important reason was that the Soviets'

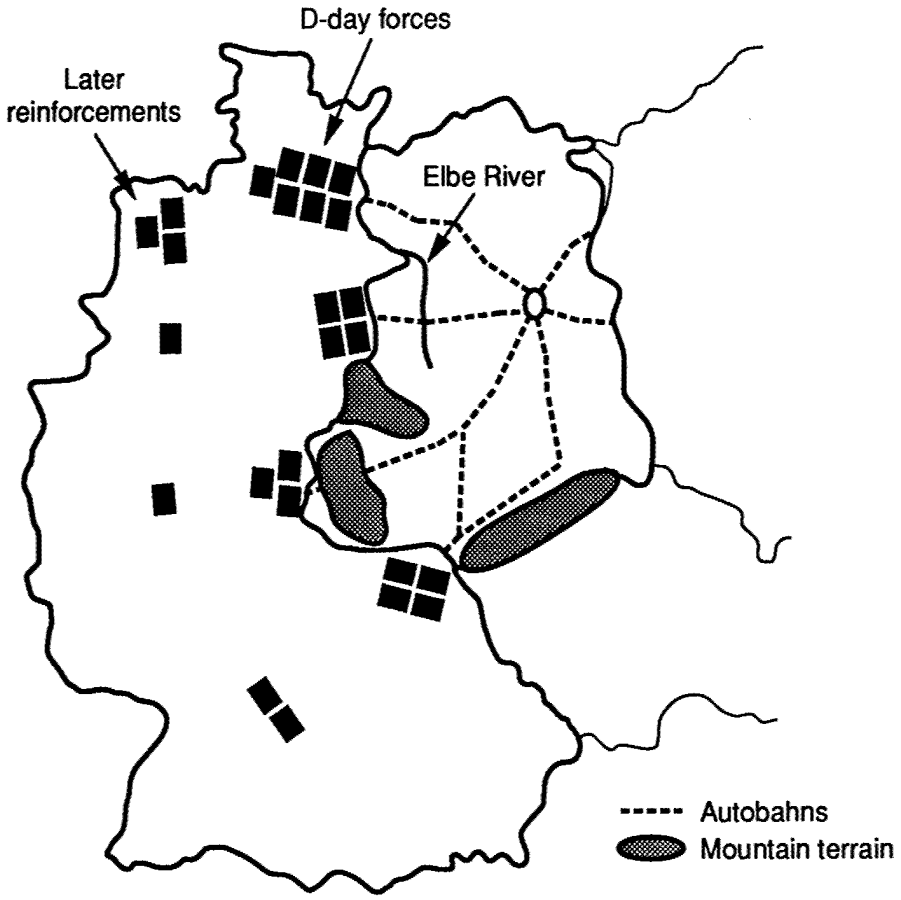

Fig. 17—A Northern-Anchored Advance into Eastern Germany

supply lines would become extended as they advanced westward while NATO's supply lines would shorten as its troops fell back on their support bases. This situation would be reversed in the new scenario, since the Soviets would be falling back and NATO would be stretching out its lines of communication. At issue would be whether NATO's logistic support forces would be able to transport adequate amounts of ammunition, fuel, and other items to support combat operations if war broke out.

Typically, Western armies design their truck transport forces and support base network to haul supplies to a distance of only about 250–300 kilometers. A Salle River defense line would be located only a short distance into eastern Germany and within range of NATO's present support bases in western Germany. An Elbe River defense, however, would be placed further eastward, especially in the area south of Berlin, where the Elbe line would lie 400 kilometers from NATO's support bases. This would place NATO's line potentially beyond the range of its ammunition and fuel supply network. A long supply line would stretch NATO's support capacity in other areas. Military police and engineers would have to service a longer road network, and heavy maintenance units would be located a greater distance

from the combat units they support. Logistic support could still be provided to an Elbe line, but at less than doctrinally prescribed levels.

This problem could be solved if NATO's ammunition dumps, equipment stocks, and support bases were transplanted into eastern Germany. Alternatively, NATO could move its truck transport units forward and rely on railroads to carry supplies to them. Both measures, however, would consume valuable time. In the interim, NATO's combat capability might be eroded by the constraints that would be placed on the expenditure of ammunition, maintenance repair, and other important functions.

The roots of NATO's military strategy are deeply implanted in the layer cake array at the old inter-German border. NATO's present posture is not designed for what inevitably would be a complicated and demanding move eastward. The further east NATO's forces were called on to advance in a crisis, the greater the stress placed on NATO's current arrangements. The transition period ahead thus is likely to be a mild forerunner of what is likely to occur once NATO's defense plans begin concentrating on the mission of protecting all of eastern Germany with a defense posture that remains anchored in western Germany.

As is the case with any new and complex military operation, the task of developing a NATO capability to move eastward will require a major effort to throw out old plans and craft new ones. In the absence of this effort, NATO's posture will remain primarily tied to the defense of western Germany. In a crisis involving eastern Germany, it would be able to move eastward only in an improvised, lurching, and inefficient way. Planning cannot solve all the problems that would arise, but it could identify the forces that would be available, determine how and where they are to move, and arrange logistic support for them. This alone could spell the difference between success and failure.

Equally important, an effort to develop revised military plans for this transition period could help NATO embark on the task of determining how it is to practice coalition warfare in the post–Cold War era ahead. Once Soviet forces are entirely withdrawn from Germany, a new European military balance and stability equation will confront NATO with the job of making even more fundamental changes in its defense plans. If nothing else, the upcoming transition period offers NATO an opportunity to cut its teeth on this important challenge.

DEFENDING EASTERN GERMANY: THE END GAME

As it sets out to design military plans for the period after Soviet forces have withdrawn, NATO will first need to address the new spatial relationships of the military balance in Central Europe. During the Cold War, NATO and Soviet forces directly confronted each other across a tense border. In the post–Cold War era, they will be separated by a distance of 1000 kilometers or more. The landmass between them will constitute what amounts to a largely demilitarized buffer zone. The Soviets, in particular, will lack an established military infrastructure there that could permit rapid transit across it. Whereas previously a war could have broken out on short notice, now it could occur only after a more prolonged period of forward movement by both sides.

As Fig. 18 shows, Soviet forces would have to mobilize within the USSR and move 600 kilometers across Poland before reaching Germany's borders. NATO's forces, which mostly will still be based in western Germany, would themselves have to move 400 kilometers just to reach the Oder-Neisse rivers. Beyond doubt this physical separation of forces will have a stabilizing effect. Both sides, but particularly NATO, no longer will have to worry about surprise attack. A buffer zone nonetheless can be crossed and potentially at a faster rate

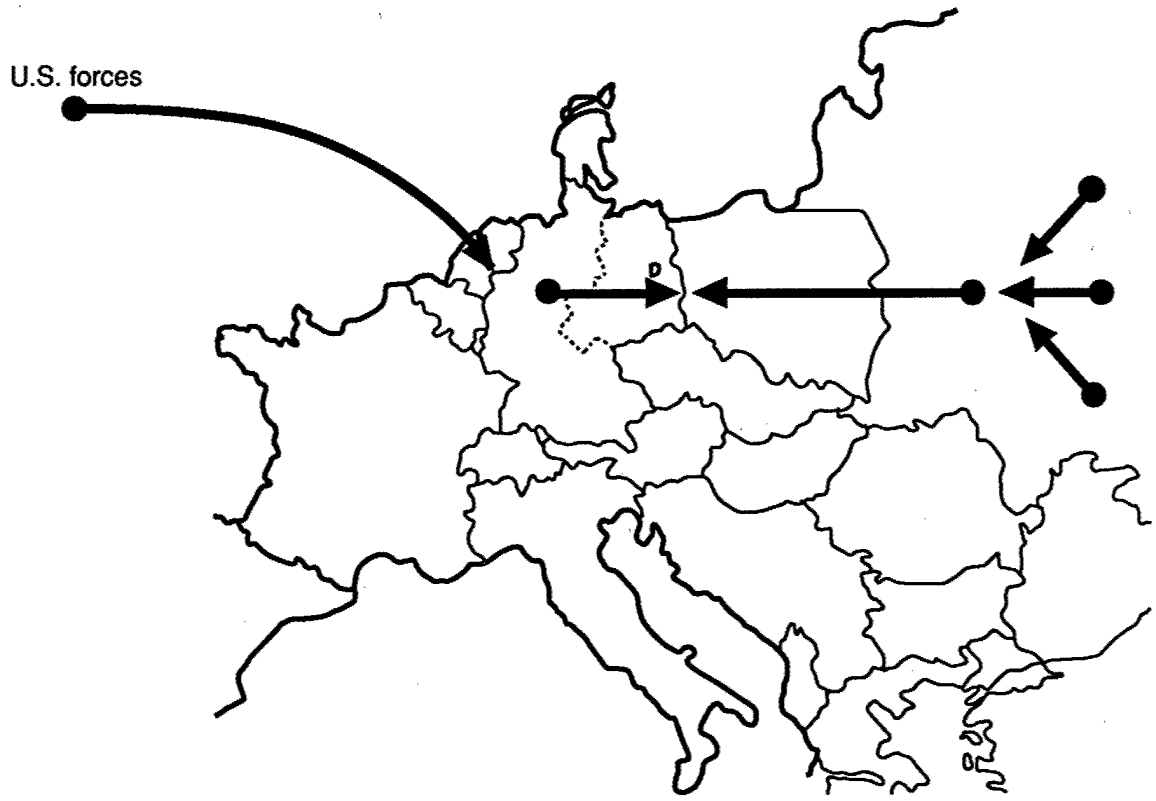

Fig. 18—The New Dynamics of Military Stability

than surface appearances might suggest. The military balance therefore will not be so inherently stable that war is ruled out solely by physical constraints on the use of force. In particular, the Soviets will still be capable of mounting an offensive campaign in Central Europe should their future political goals lead them in this direction. NATO will need to turn its defense planning efforts to address this new form of risk.

Although the risks ahead will be lower than in past years, NATO will continue to confront a force generation problem in Central Europe. In some respects, this problem will be less formidable because Soviet forces would have to move longer distances and across a Polish nation that might not welcome their presence. In other respects, however, it will pose unique and potentially troublesome challenges of its own. The central phenomenon here is not that the Soviets alone would require more time to launch an attack, but that both sides would face constraints on projecting their forces forward. To the extent NATO benefits from this situation, it will be in relative terms.

For three reasons, NATO will need to be concerned about its capacity to stay on top of what could become a mobilization and reinforcement race in slower motion than before. First, the political dynamics of any future Central European crisis almost certainly would be more complicated and ambiguous than any Cold War scenario. Uncertainty and a natural desire to avoid acting provocatively could inhibit NATO from making decisions in a timely

fashion. Second, the task that NATO would face in moving its own forces into eastern Germany, while less daunting than that facing the Soviets, would itself not be trivial. Above all, it could not be accomplished overnight. Finally, fewer U.S. and British forces will be based on the European continent in peacetime. NATO therefore will be more dependent on the time-consuming task of bringing in outside reinforcements. For all these reasons, NATO might be best advised to view its future force generation problem not as vastly easier than before, but rather as quite different.

Solving this problem will require a concerted NATO planning effort, aimed not at girding for war but at insuring against an improbable but dangerous event. This planning enterprise, however, cannot afford to focus only on the force generation question. An equally important issue is NATO's future theater employment doctrine, including the twin tasks of fashioning an intelligent force array and assigning national responsibilities for implementing it. If a linear defense of eastern Germany would not make sense, then what should be substituted? And if the old layer cake no longer suffices for determining national missions, how should NATO plan to wage coalition war in the future? These issues cannot be addressed in isolation. Force generation and theater employment doctrine will be intertwined and must be addressed in tandem.

Force Generation

In the years ahead, Soviet diplomacy is unlikely to be motivated by the expansionist designs that guided it during the Cold War. Barring a wholesale reversal of Gorbachev's reforms, ideology will be replaced by a continued trend toward a pragmatic stance marked by the aim of establishing stable, economically profitable relations with the West. Especially if the USSR's political system becomes increasingly pluralist and fragmented, NATO will not need to worry about a crisis that has its origins in naked and premeditated Soviet aggression.

A far more realistic worry is a crisis growing out of Central Europe's tangled politics coupled with the USSR's perception of its own national interests and tenuous internal situation. Precisely how such a crisis could lead to the extreme step of Soviet aggression against Poland and then Germany is difficult to foresee. Clearly it would have to grow out of a different political psychology than exists today and evolve over a long period of declining relations. In particular, the emergence of tense Soviet-German relations would be necessary. Emotional currents, however, can change in ways that give rise to tensions, minor conflicts, and other trends that lead otherwise peaceful nations to begin thinking about using military force against each other. In this climate, a crisis conceivably might begin in the Balkans or the Middle East and then spread northward in a way that defies understanding and rational priorities. As 1914 showed, escalating spirals of this sort can happen even if beforehand they seem almost beyond the limits of plausibility.

This scenario is only one among a variety of possibilities, not all of which would involve Central Europe. It is equally plausible to imagine NATO squaring off against the Soviet Union in other places. Turkey and Norway are two possibilities. Questions about its antecedents and its low likelihood aside, a crisis culminating in a Soviet attack on Poland and Germany nonetheless is a scenario that would strike at the heart of the whole European security order. For that reason, it will remain central to NATO's efforts to define the kind of military insurance policy that should be retained for all of Europe. NATO's capacity to defend Germany's borders (and the borders of other members) will be an important test of military sufficiency for NATO's defense posture. A transparent NATO capacity to defend

Germany could indirectly help deter any Soviet aggression against Poland or other East European nations, a task that also is likely to begin entering NATO's security calculus in the years ahead.

The prospect of employing a scenario that starts elsewhere and then leads to a Central European war confronts NATO's planning process with a force generation problem wholly unlike what the Cold War offered. In particular, NATO would face even greater uncertainties than before on the important issue of how fast Soviet forces could arrive in strength at Germany's eastern border. Figure 19 helps illuminate these uncertainties by displaying alternative timelines for the Soviet buildup rate in this scenario.

As this figure suggests, the Soviets conceivably could generate a 65-division force at the German border and attack in as little as 40 days. This could be the case if the task of physically moving these forces across the Soviet and Polish rail network was the only constraint facing the Soviets. Even in the absence of active NATO interference, however, three constraints could intervene to drag out the buildup process considerably. First, the need to conduct refresher training for cadred Soviet units maintained at low readiness could add at least three weeks, and possibly six. Second, a delay of two to four weeks could arise from the requirement to deploy large stocks of ammunition and fuel and to establish secure infrastructure facilities (e.g., airbases, command posts, and ammo dumps) before attacking. Third, minor Polish resistance could prevent the Soviets from attacking Germany by a week or so;

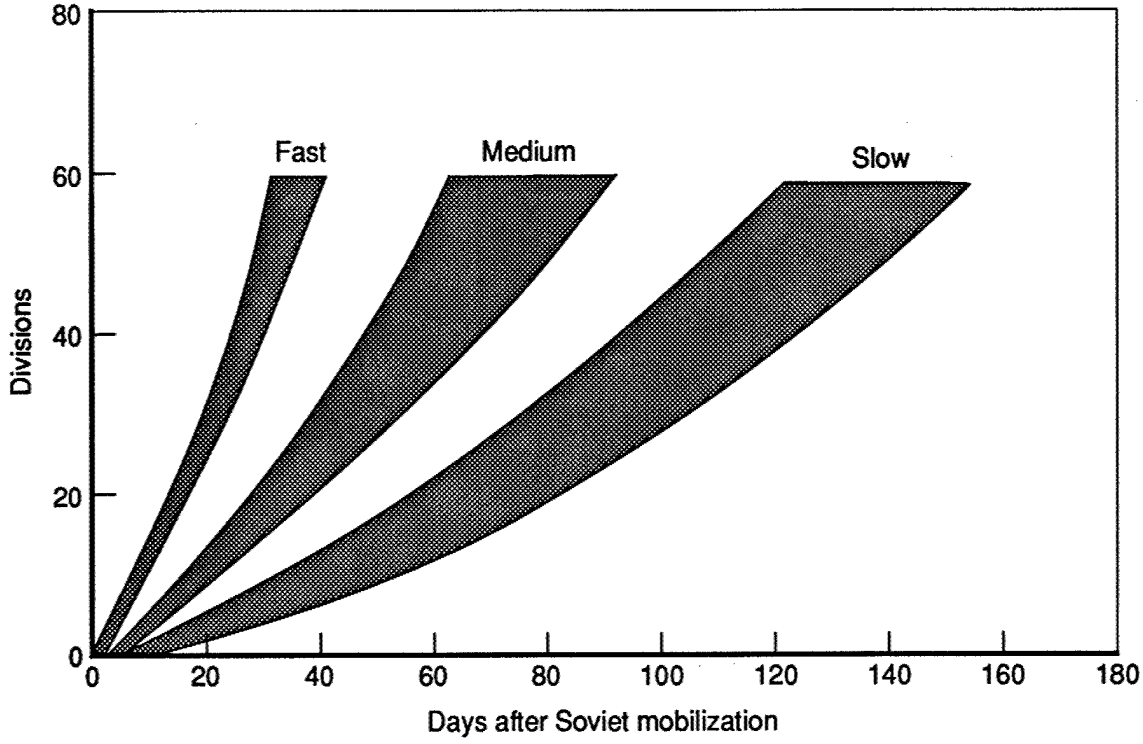

Fig. 19—Alternative Soviet Buildup Rates in Central Europe

major opposition, requiring the Soviets first to overpower the Polish Army, could drag out the buildup process by a month or even longer.

The total delay would depend upon the exact effects of these constraints and upon how the Soviets reacted to them. The task of estimating future Soviet buildup rates therefore is confronted with major uncertainty, and the subject deserves far greater treatment than space allows here. What can be done here is to bound the range of uncertainty. If the minimum time is 40 days, a period of about 120–150 days represents an outer limit for how long it might take the Soviets to overcome powerful constraints in all three areas. A period of 60–90 days represents a middle-ground estimate that allows for a substantial slowing of the Soviet effort, but not paralysis.

Taken together, these factors suggest that NATO will have a good deal more time than in earlier years to conduct its own force generation effort. This extra time would help ease the planning dilemmas that NATO defense programs traditionally have faced in Central Europe. The inherent uncertainty about Soviet capabilities, however, creates powerful reasons why NATO could not afford the luxury of approaching the mobilization and reinforcment process in a leisurely fashion. Stated simply, a NATO force generation plan pegged to a late D-Day could get caught short if things went better for the Soviets, or worse for NATO, than it assumes.

As a result of this wide range of uncertainty, NATO will face a complicated choice in deciding to prepare for the best, the worst, or somewhere in between. In the actual event, NATO might well have the benefit of four months or more of warning time between a Soviet mobilization and an attack on Germany. But a NATO decision to assume this will be the case might be an imprudent bet. At the other end of the spectrum, a decision to plan against a full-scale attack within 40 days would improbably calculate that the Soviets could race unimpeded through Poland and then attack Germany without pausing to regroup. Assuming that realism and prudent conservatism are the guiding criteria, a NATO decision to plan on the basis of D-Day's occurring at Soviet M+60 or somewhat later would represent a middle-ground choice.

The likelihood that the Soviets could not attack Germany in strength before M+60 does not guarantee that NATO would have two months at its disposal to prepare its defenses. A second critical factor in this equation would be NATO's ability to react to warning of an impending attack by itself deciding to mobilize. Figure 20, which assumes that D-Day occurs at Soviet M+60, illuminates some of the dilemmas that would confront NATO as the Soviet buildup process unfolds.

Within the first week or so, NATO would receive unambiguous indicators that the Soviets had begun a full-scale mobilization. Along with these indicators would come increasingly obvious violations of the CSCE and CFE accords. These provocative actions, however, would not necessarily be confirmation that the Soviet government had firmly decided to embark on aggression in Central Europe. Ambiguity would remain not about the USSR's growing military capacity, but about the political intentions of its government. This ambiguity would begin clearing by M+30, once Soviet forces had begun entering Poland. But even then, uncertainty would remain about whether its aggressive designs extended to Germany. This uncertainty would be resolved once Soviet forces began approaching the German border, which would not occur until M+40 or thereafter, only three weeks before D-Day.

The ambiguous nature of the political-military situation early in the Soviet buildup process would confront NATO with agonizing dilemmas. Prudence would call for a NATO decision to begin mobilizing very early in this cycle. The need to deter the Soviets and to show

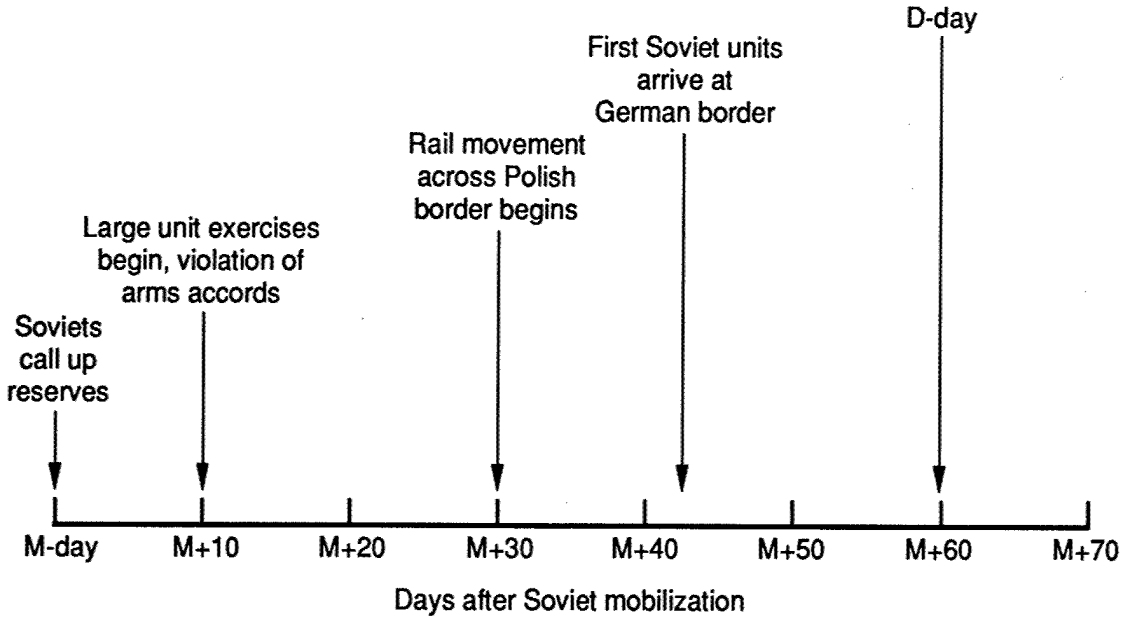

Fig. 20—Military Scenario for a Soviet Attack on Germany (late 1990s)

resolve would further reinforce the incentives for taking this step. Especially if NATO's own capacity to respond rapidly was not impressive, military exigencies would also drive NATO in this direction. At the same time, lingering uncertainty about the USSR's political intentions could have an inhibiting effect. Reinforcing NATO's hesitancy might also be a natural desire to avoid either worsening the situation through provocative NATO actions or giving the Soviets a pretext to invade. To the extent that these concerns prevailed, NATO's decision to mobilize and reinforce plausibly might be delayed by several days or even weeks. The later the NATO decision to mobilize, reinforce, and move eastward, the less time would be available to prepare NATO's defenses. This delay, in turn, could substantially shorten the time that NATO would have available to generate an adequate defense force in eastern Germany.

The task facing NATO in deploying a large posture in eastern Germany would be more complex than simply deploying forces to the old inter-German border. NATO is currently capable of erecting its layer cake within three to four days, but because most NATO forces will remain at their present locations, more time would be needed to build similar strength in eastern Germany, especially along the Oder-Neisse rivers. The peacetime presence of some 50,000–100,000 active German soldiers there would enable the German government to quickly establish a ground and air screen along the Oder-Neisse rivers and preclude any early attack by highly ready Soviet forces (e.g., an airborne assault on Berlin). NATO units based in western Germany, however, would arrive well later. The process could be hastened if the German government were to use the years ahead to ensure that an adequate military infrastructure is maintained in eastern Germany. This would obviate any need to undertake such complex measures as preparing command posts, airbases, and supply depots. But even so, an extra week or two could be consumed in moving NATO ground combat forces, support units, and air wings into eastern Germany. Reinforcement from the United States would

confront similar problems. Like the possibility of a late NATO decision to mobilize, this delay also would cut into NATO's margin of safety.

For these reasons, a buildup race in Central Europe would be characterized by a NATO effort to assemble an adequate defense posture in eastern Germany before the Soviets could deploy enough forces to cross the Oder-Neisse border. Military stability, in turn, would be measured by NATO's capacity to win this race. To the extent that NATO fell short of being capable of performing this mission in a timely fashion, the situation would be unstable.

NATO's ability to win this race could be affected by its ability to come to grips with the painful decision to mobilize. The following two figures illustrate why. Figure 21 measures how NATO's buildup rate would match its force needs in eastern Germany if NATO were to mobilize almost immediately after the Soviets began mobilizing. NATO would need first to establish an Oder-Neisse defensive screen of 10–20 DEFs and later to generate a total of 33–48 DEFs. The range of future capabilities is a function of whether NATO's present posture is maintained or reduced by 25 percent in size.

Provided NATO decides to mobilize immediately, its time-phased buildup rate will remain adequate for meeting its requirements. This situation would still prevail even if NATO's forces are, as is possible, reduced by 25 percent. Indeed, NATO probably could afford to reduce its readiness levels and still be able to respond effectively. As Fig. 22 shows, however, a less sanguine picture would emerge if NATO were to delay its decision to mobilize by three weeks or more. In that case, the situation would be touch-and-go throughout. If the Soviets were to attack as early as M+60, NATO's posture would not be well-insured, especially if it is cut by 25 percent or more. Little room would be available for reducing NATO's

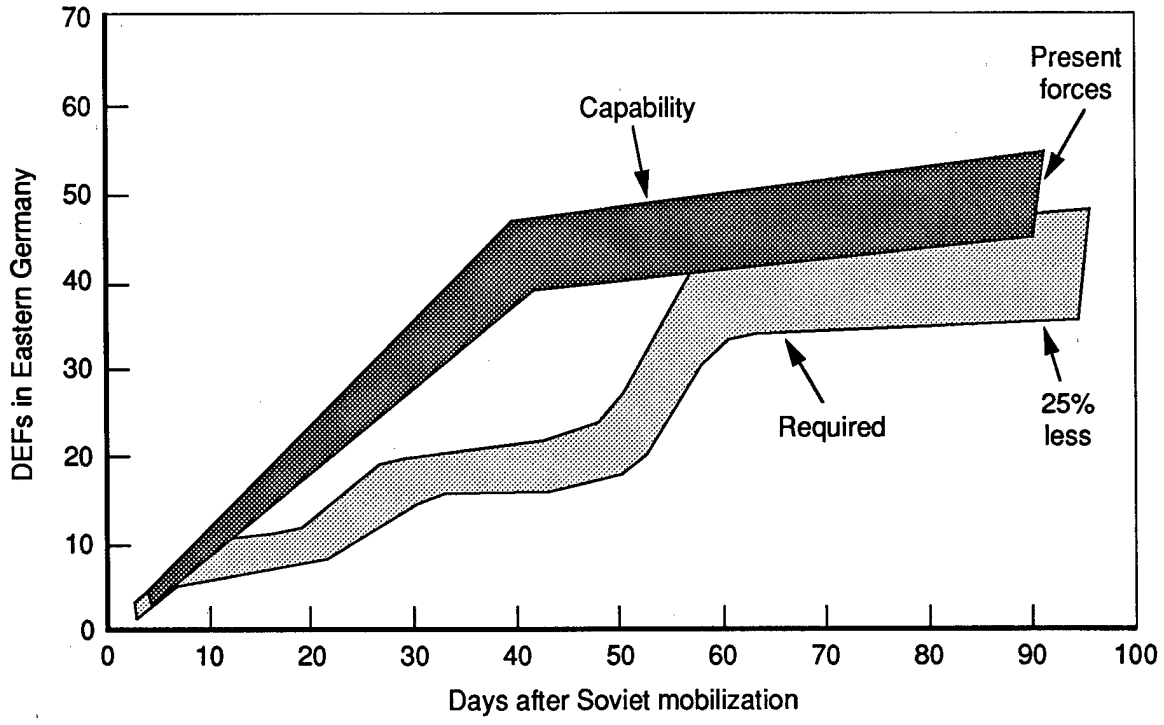

Fig. 21—Adequacy of NATO's Ground Buildup Rate: Early Decision to Mobilize (Late 1990s)

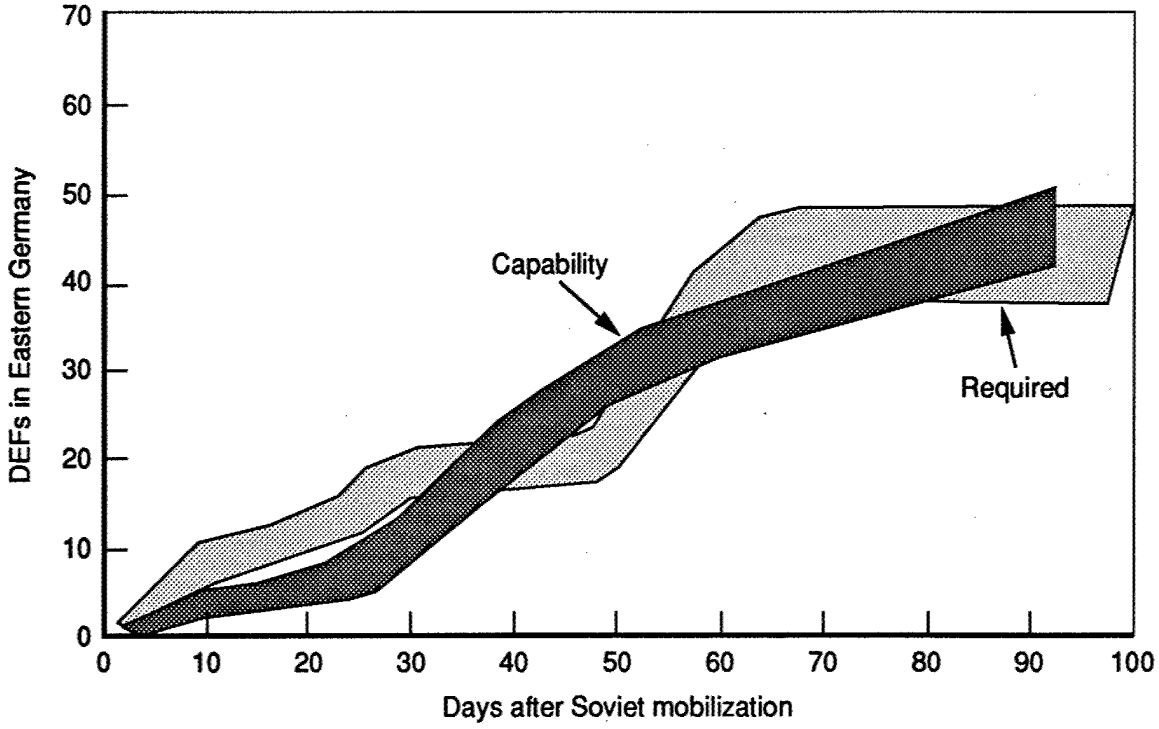

Fig. 22—Adequacy of NATO's Buildup Rate: Late Decision to Mobilize (Late 1990s)

readiness. In fact an incentive would exist for increasing the readiness and deployment rates of later-arriving units, especially speeding up the deployment of U.S.-based forces in CONUS. Any decision to delay mobilization beyond this three-week period could result in a NATO buildup rate that is well below needed force levels.

A similar but less stressful situation applies to NATO's tactical air forces. Figure 23 displays how NATO's tacair capabilities compare with estimated requirements. The requirement band is based on the assumption that NATO needs sufficient forces early to provide an air defense screen over Western Europe and enough forces later to keep the air balance close to parity. NATO's ability to generate adequate air forces is less sensitive to the decision to mobilize because NATO air forces in Central Europe can move forward more quickly than can ground units, and U.S. air forces based in the United States can deploy overseas quite quickly.

Estimates of NATO's future required force levels, readiness standards, and decision-making propensities are interconnected. The transformation of the military balance in the years ahead will give NATO more time to prepare while allowing it to cut its force levels and readiness standards along the lines now being contemplated by member nations. If NATO decides to scale back in all three areas, however, the point could be reached where its forces and buildup rate would not be adequate even for the less threatening era ahead.

NATO's ability to react to ambiguous warning is an especially important factor in this calculus. Even taking into account the need to move large forces into eastern Germany, NATO apparently will have greater time at its disposal to read the situation and then to

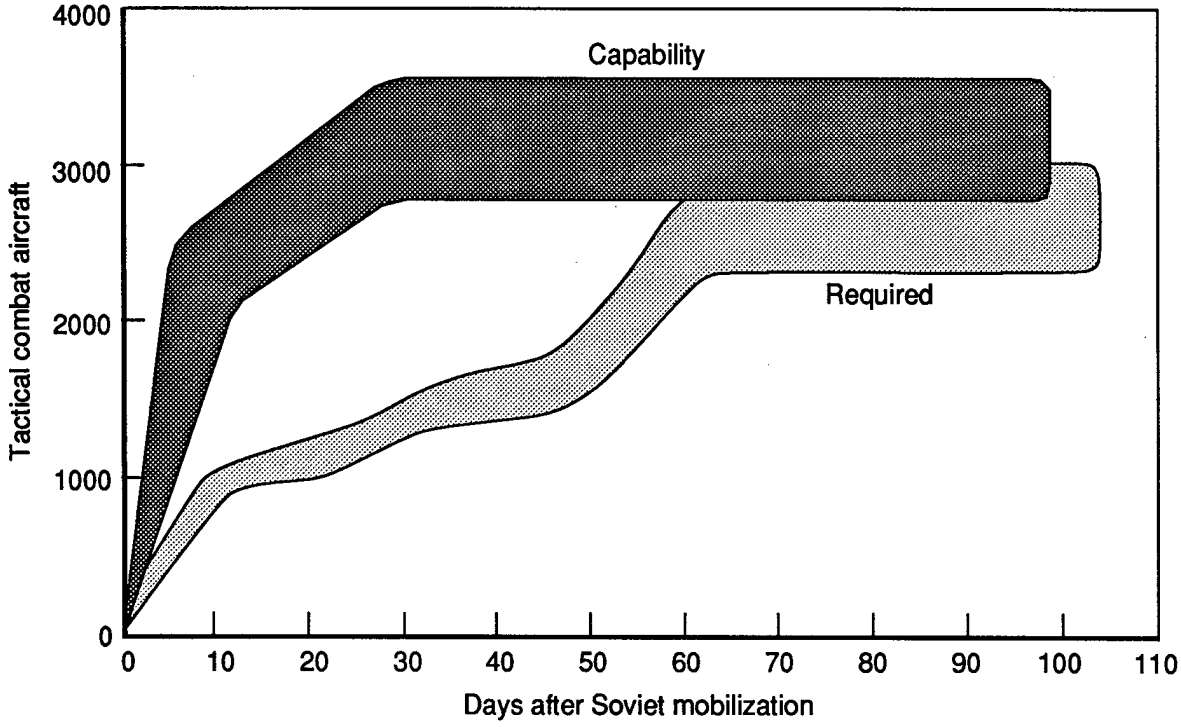

Fig. 23—NATO's Tactical Air Forces: Capabilities vs. Requirements (Late 1990s)

mobilize effectively. But this larger time window is not unlimited, nor is it without drawbacks of its own. NATO might find itself compelled to plan on mobilizing while Soviet forces are still deployed in the USSR, before they have crossed into Eastern Europe. A slower Soviet buildup rate, of course, could ease this pressure to decide prematurely. In particular, the Soviets might encounter time-consuming problems crossing Poland. The problem facing the Alliance's defense plans is that NATO has no way of knowing in advance how the Polish government might respond to a Soviet intrusion. Indeed, NATO might not know the answer to this important question even at the onset of a crisis. Moreover, for reasons of its own security, NATO could not afford to remain indifferent to Poland's fate even though that nation is not a member of the Alliance. An early NATO decision to mobilize might help deter a Soviet advance into Poland, a consideration that would have to be taken into account in calculating NATO's willingness to respond.

NATO will need to remain sufficiently alert to be able to reach a mobilization decision more quickly than surface appearances might suggest is necessary. Also, NATO will need to retain a sufficiently large and ready military posture capable of generating forces at the required rate. Decisions on NATO's posture will have to take into account the need to move longer distances than before and the possibility of undesirable delays in the decision to mobilize. As matters now stand, present member nation plans for reducing NATO's force size and readiness levels appear to fall within the boundaries of safety. But considerably deeper cuts in either size or readiness could unhinge NATO's military strategy. In any event, these are tentative conclusions. Much remains to be learned about the military regime ahead, and NATO's force generation rate will need to be the subject of extensive review and careful study before it can reach definitive conclusions.

Criteria for Evaluation

As it shapes its military plans for protecting eastern Germany, NATO will need to design a new force employment concept that can replace the old linear defense and layer cake array. The following criteria stand out as important in guiding the selection process.

- The new doctrine must be militarily effective, thereby helping to achieve the security goals of deterrence, defense, and military stability. It must reduce the vulnerabilities that a continued reliance on linear defense at low force levels would create. It must help NATO compensate for continuing to be outnumbered by at least some margin. It must be executable with the forces that realistically will be available and thereby facilitate NATO's efforts to maintain a stalwart conventional defense capability.
- The new doctrine must fit within the framework of the mobilization and reinforcement dynamics that could unfold in a future Central European crisis. In particular, it must not depend on an early NATO decision to mobilize or propel the Alliance into mobilizing prematurely in a crisis. Also, it must provide a flexible basis for making adjustments if NATO's mobilization decision comes later than is desirable.
- It must provide confidence in its ability to protect Germany's eastern territories insofar as forward defense is militarily achievable on the modern battlefield. Its execution should not depend heavily on trading space for time in some major way.
- It should be consistent with NATO's goal of making nuclear escalation a tactic of last resort. But it also should provide a basis for escalating should that step become necessary.
- It should provide a sound framework for assigning national missions, allocating military burdens, and adjusting command relationships in the post–Cold War era. In particular, it should call on each nation's military forces to contribute in the ways for which they are best suited. It should encourage the Alliance to pursue greater specialization and division of labor in areas where this step would enhance NATO's security.
- It should help promote NATO's unity. It should avoid creating sharp conflicts among NATO's partners and help foster support for the Alliance's strategy among the domestic constituencies of member nations. It should provide Germany a continuing rationale for remaining firmly within NATO's fold and encourage France to draw closer to the integrated military command. It should maintain a strong role for Britain and encourage the Low Countries to continue contributing to NATO's defenses. Also, it should help preserve a close transatlantic relationship, providing an important role for U.S. forces to play, thereby underscoring the continued existence of extended nuclear deterrence and avoiding the appearance of decoupling.
- It should enhance NATO's coalition strength. It should help NATO's members to pursue greater integration in such areas as multinational combat forces, combined logistics, and armaments cooperation. It should provide a basis for combined planning and encourage NATO's members to seek solutions to the problems of establishing a coalition defense posture. Also, it should encourage greater cooperation among the West European nations in appropriate areas and be consistent with their efforts to develop the European Community and the West European Union (WEU).
- It must be financially affordable. In addition to being executable with the force levels that will be available, it also must be broadly consistent with future U.S. and

allied programs for modernization, readiness, and sustainability. This does not imply that NATO's future employment doctrine should refrain from calling for vigorous defense programs. Ideally it should help prod NATO's members into maintaining adequate defense efforts. But in the process of doing so, it should establish a compatible relationship between NATO's future force goals and resource guidance.

- It should be fully consistent with NATO's arms control policies and help dampen incentives for a post–Cold War arms race. Also, it should reinforce the Alliance's broader efforts to promote European stability by projecting a defensive, reassuring image to the Soviet Union and the East European nations. It should be compatible with future efforts that might be undertaken to expand CSCE's role.

By virtue of their sheer numbers and the requirements they levy, these clearly are demanding criteria. In all likelihood, no single employment doctrine will prove able to satisfy all of them, especially in a maximum way. Tradeoffs, compromises, and shortfalls in some areas inevitably will be involved. However, the doctrine ultimately adopted should be able to satisfy the most important military and political criteria, while at least not doing serious damage to the rest. Within this framework, what matters most is the new doctrine's overall political and military effect.

Fortunately, the terrain in eastern Germany favors NATO in ways that make conventional defense a viable goal. The terrain north of Berlin is dominated by the wide Oder River and by marshes and waterways to the rear. South of Berlin, the Neisse, Spree, Elbe, Mulde, and Salle rivers form a succession of barriers that would be hard for the Soviets to cross. But even the task of defending favorable terrain requires some minimum level of forces and a sound employment doctrine. Other things being equal, small forces especially require a good doctrine.

NATO's ability to craft a new employment doctrine will be affected by the provision that its forces remain based in western Germany. Rear basing will not, in itself, prevent NATO from pursuing the doctrine of its choice. But NATO would require sufficient time to move its forces to their wartime locations. To the extent time is at a premium, these movement dynamics might constrain NATO's capacity to deploy its forces in a crisis. A delayed deployment, in turn, could shape NATO's doctrine for employing these forces. Time might not be available to reach forward positions and develop the prepared positions that are the anchor of a frontal defense. The result would be to push NATO in the direction of a more fluid doctrine that places less emphasis on defending fixed terrain features and more on coordinated battlefield maneuvers.

Reinforcing this trend will be the inability of most NATO forces to conduct peacetime training exercises in eastern Germany. German military personnel will have access to the forward areas, but non-German forces will not be able to gain first-hand experience with the terrain that they might be required to defend. Since knowledge of the terrain is critical to a defender's ability to execute a positional defense, this factor also increases the need for these forces to adopt a more fluid form of warfare.

NATO's ability to tilt in the direction of maneuver warfare will be constrained by political considerations, among them the emphasis being given to reassuring Poland. The London Summit's commitment to nonaggression is one example of NATO's ongoing efforts to signal the Polish government that Germany's unification will not produce any security threat to it. Chancellor Kohl's reaffirmation and the subsequent German-Polish Treaty enshrining the current borders reinforce this commitment. Concern for Poland's sensitivities is one

additional factor behind NATO's willingness to refrain from deploying its forces into eastern Germany. In the years ahead, this concern will probably be translated into official reluctance to endorse any employment doctrine that allows for defensive NATO operations in Poland. If so, this dictate will complicate NATO's freedom of choice in selecting an alternative to linear defense.

NATO's employment doctrine will not be able to take formal advantage of the territorial buffer that Poland offers. In particular, NATO will be unable to anchor itself on plans to move forward into that country and to conduct the kind of free-wheeling, space-trading mobile warfare that otherwise might accompany a wholesale shift away from linear defense. This situation might change if Poland were to issue an appeal for help in an actual crisis. For this reason, development of appropriate contingency plans makes sense, along with prudent measures to help Poland defend itself. There is of course a critical difference between contingency plans and the programmatic plans that officially shape NATO's force posture and employment doctrine. In the years ahead, Poland's stance seems likely to remain a wild card. Out of concern for not offending Poland and for ensuring that NATO's doctrine is executable, NATO's programmatic plans will need to be anchored on the assumption that combat operations would begin on Germany's borders rather than points eastward.

An employment doctrine that conducts NATO's defense efforts from Germany's borders westward is not antithetical to defensive maneuver warfare, but it could compel NATO to undertake these operations on its own territory and possibly in its rear areas. Germany's unification will provide NATO greater depth for operational maneuver and space-trading than before. Along the Oder-Neisse rivers, moreover, NATO will be operating behind a straight line, not the concave line that made defense difficult on the inter-German border. These changes will open up the battlefield for mobile warfare and nonlinear tactics in ways that were not possible in past years.

NATO's capacity to move in this direction inevitably will be complicated by the presence of Berlin, Dresden, and other urban areas near the Oder-Neisse rivers. NATO may be able to part company with an inflexible commitment to linear defense, but whether it will be able to cast aside forward defense is another matter. Without a decision by the German government to endorse other priorities, concern for this goal is likely to play an important role in NATO's criteria for designing a new employment doctrine.

The years ahead are likely to see the emergence of several doctrinal candidates, each with its own assets and liabilities. Two options in particular meet these criteria reasonably well. Both would replace NATO's linear defense and layer cake with an "echeloned" defense that provides frontal strength and greater depth in the rear areas. Two ways of implementing each of these parent doctrines allocate national missions according to differing priorities. One would push NATO in the direction of multinational forces and continued emphasis on a layer cake made up of units from several nations, and the other would encourage greater national specialization by assigning the frontal defense mission primarily to German forces and the reserve mission to the other nations.

Soviet Operational Concepts

An analysis of NATO employment doctrine must begin with an appraisal of likely Soviet operational concepts. Figure 24 illustrates how a three-front Soviet attack with 65 divisions and 2800 tactical combat aircraft might be conducted. It assumes that the Soviet goal would be to overrun at least eastern Germany and to defeat NATO's forces there. The eastern

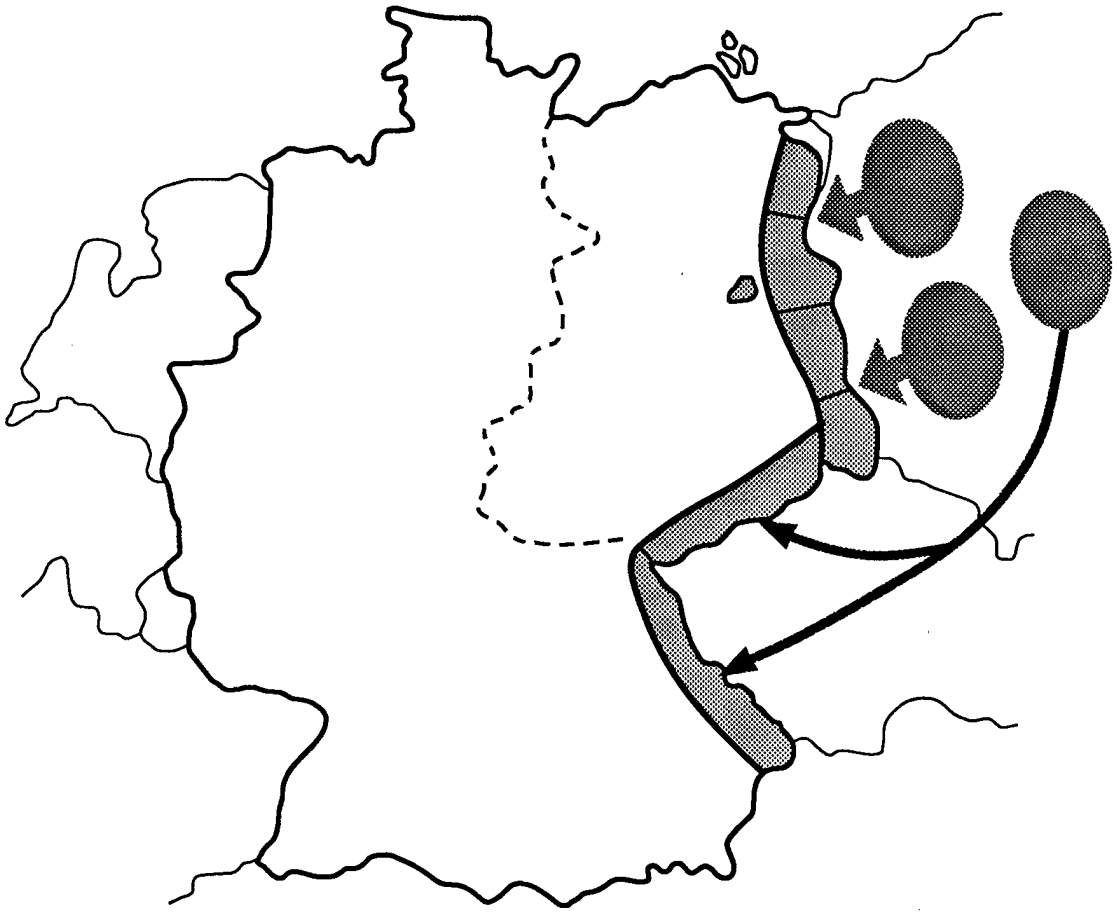

Fig. 24—Soviet Operational Concepts

German border opposite Poland is only about 300 kilometers long. Therefore, only two of the three Soviet fronts could be committed in the initial assault echelon. The third front would have to function as a second echelon reserve that could provide extra staying power for the committed forces or be used to prosecute a breakthrough and envelopment campaign. Alternatively, this third front could swing south through Czechoslovakia and either drive into Bavaria or, more likely, try to outflank NATO's forces to the north.

This Soviet attack, with its ability to conduct a powerful frontal assault or to engage in a flanking maneuver, would pose a serious threat to eastern Germany and NATO's forces there. However, it would be confronted with terrain impediments that, at a minimum, would hamper its effectiveness and slow down its advance. Soviet forces attacking directly into eastern Germany would have to cross the Oder-Neisse rivers; particularly in the north, the Oder river is a major barrier. Once across, Soviet forces operating north of Berlin would confront swamps and marshes that could slow their advance. In the center, Berlin itself is a major barrier to a quick Soviet drive, especially since Germany's autobahns there lead directly into Berlin. To the south, Soviet forces would have to cross the Elbe, Spree, and Mulde rivers, a sequence of barriers that would have a retarding effect.

Any Soviet operations in Czechoslovakia would confront similar problems. The rough terrain and the Labe River would inhibit Soviet forces from moving quickly. Equally important, the Erzgebirge Mountain range, which guards the border between eastern Germany and Czechoslovakia, would complicate any Soviet drive aimed at striking NATO's forces from their southern flank. Soviet forces transiting this area would need to pass through narrow, defensible defiles. Only the area on the eastern corner of the German-Czech border is sufficiently open and flat to permit normal offensive operations. A Soviet attack on Bavaria would also confront difficult terrain. In addition, it would extend Soviet supply lines and run the risk of diverting large forces to secondary objectives. Even if the Czech Army did not mount serious opposition and NATO itself refrained from interfering, an advance through Czechoslovakia would hardly be inviting.

A Soviet attack on eastern Germany would not be an easy proposition. A flanking maneuver through Czechoslovakia might be attempted to divert NATO's forces from the Oder-Neisse border region, but such an operation would itself not be easy to undertake. To the extent this operation could be discounted or at least brought under control, NATO's forces could focus their efforts on defending eastern Germany. There they would benefit from advantageous terrain, and they would be defending along a straight front. No longer would they have to labor under the disadvantage of a concave line. Since the Soviets would not enjoy interior lines of communication, NATO's forces would be in a better position to concentrate and counter-concentrate with equal rapidity.

Eastern Germany thus is defensible; the difficulties that the Russian Army encountered in taking Berlin in 1945 are testimony to this fact. NATO can aspire to conduct a stalwart conventional defense there. It does need adequate forces to take full advantage of the favorable terrain, which would pose no insurmountable barriers to the Soviets if they were to face only light NATO opposition. And NATO also needs a well-designed employment concept that will help it checkmate the range of offensive operations that the Soviets could mount.

Defense in Depth

A first candidate for performing this checkmating function is an employment doctrine of "defense in depth," which is displayed in Fig. 25. This doctrine would take advantage of the greater strategic depth that unification will bring by keeping NATO's center of gravity well to the rear. A light screening force would be deployed along Germany's borders, including the Oder-Neisse rivers. Deployed some 150 kilometers to the rear would be two to three mobile NATO corps; farther back, in western Germany, would be NATO's remaining forces. In this scheme, only about 25 percent of NATO's total posture would enter eastern Germany itself. The remainder would plan to fight in the vicinity of the old inter-German border.

This employment doctrine has three attractive features. First, it would alleviate the risks associated with a linear defense. NATO's forces no longer would present a forward-anchored and brittle frontal array. The Soviets would have to fight their way through three successive layers of NATO defenses, the strongest element of which would be the final layer. As a result, the Soviets would be denied any prospect of a quick and easy victory through a maneuver strategy. Second, this employment doctrine would enable NATO to adopt a war-fighting concept that conforms to the Kohl-Gorbachev agreement. Eastern Germany could be defended by German forces alone, while NATO forces from other nations would remain on western German territory. Third, this doctrine would have the least disruptive effect (of all the concepts studied here) on NATO's logistic system and infrastructure. Most NATO forces

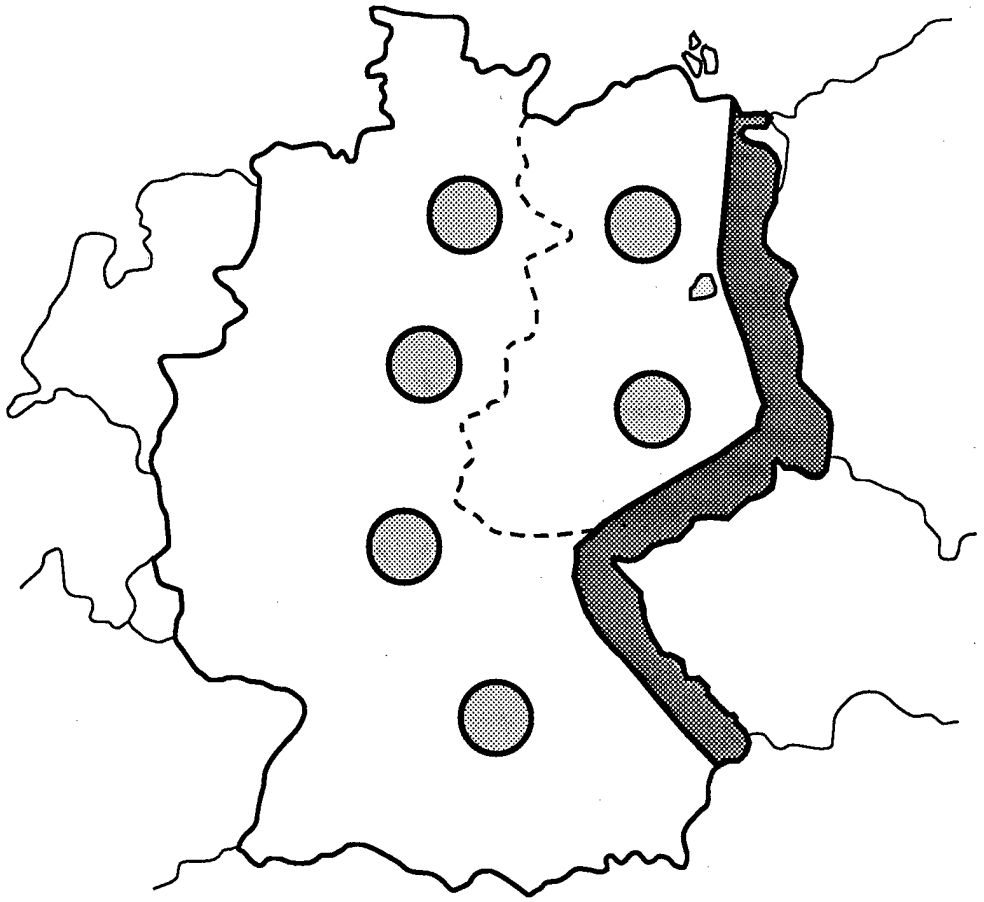

Fig. 25—Defense in Depth Has Problems

would continue to operate within the immediate vicinity of present supply depots and support bases.

This doctrine is well suited to NATO's defense requirements in the short run. Whether it is a suitable long-term solution is another matter. One drawback is that this doctrine would not do a good job of forward defense; Berlin, Dresden, and other German urban areas would be left exposed to quick seizure by advancing enemy forces. This doctrine would also physically separate NATO's forces in ways that might have adverse operational effects. In particular, the corps formations in eastern Germany would fight on their own and would not draw support from the forces deployed further to the west. Outnumbered by about 5:1, these formations might be overpowered, thereby enabling the Soviets to capture all of eastern Germany. Equally worrisome, NATO's entire posture might be so fractured into separate parts that it is left vulnerable to defeat in detail. The NATO forces deployed in western Germany might be too small to contain a follow-on enemy attack after the battle for eastern Germany had ended. Defense in depth thus solves the problem of a vulnerable linear array, but it might leave an entirely different one in its wake.

Mobile Defense

A doctrine that would endeavor to solve the problem of defeat in detail is that of mobile defense. As Fig. 26 suggests, this doctrine would mass NATO's forces together. NATO's posture would endeavor to defend eastern Germany by waging mobile warfare through meeting engagements. For example, NATO might employ classic "hammer and anvil" tactics aimed at blocking enemy armored columns and then launch attacks against the enemy's flanks. The goal would be to chop up enemy formations and then defeat them in detail. In essence, NATO would adopt a maneuver strategy of its own.

In addition to massing NATO's forces together, this doctrine has the attractive feature of employing NATO's modern ground and air equipment to its fullest advantage. As a result, Soviet forces would confront the risk that aggression might result in a decisive operational defeat to themselves brought about by sweeping NATO battlefield maneuvers. The drawbacks to this employment doctrine, however, are threefold. First, at least temporarily, a mobile defense might cede Berlin and other urban areas near Germany's eastern border.

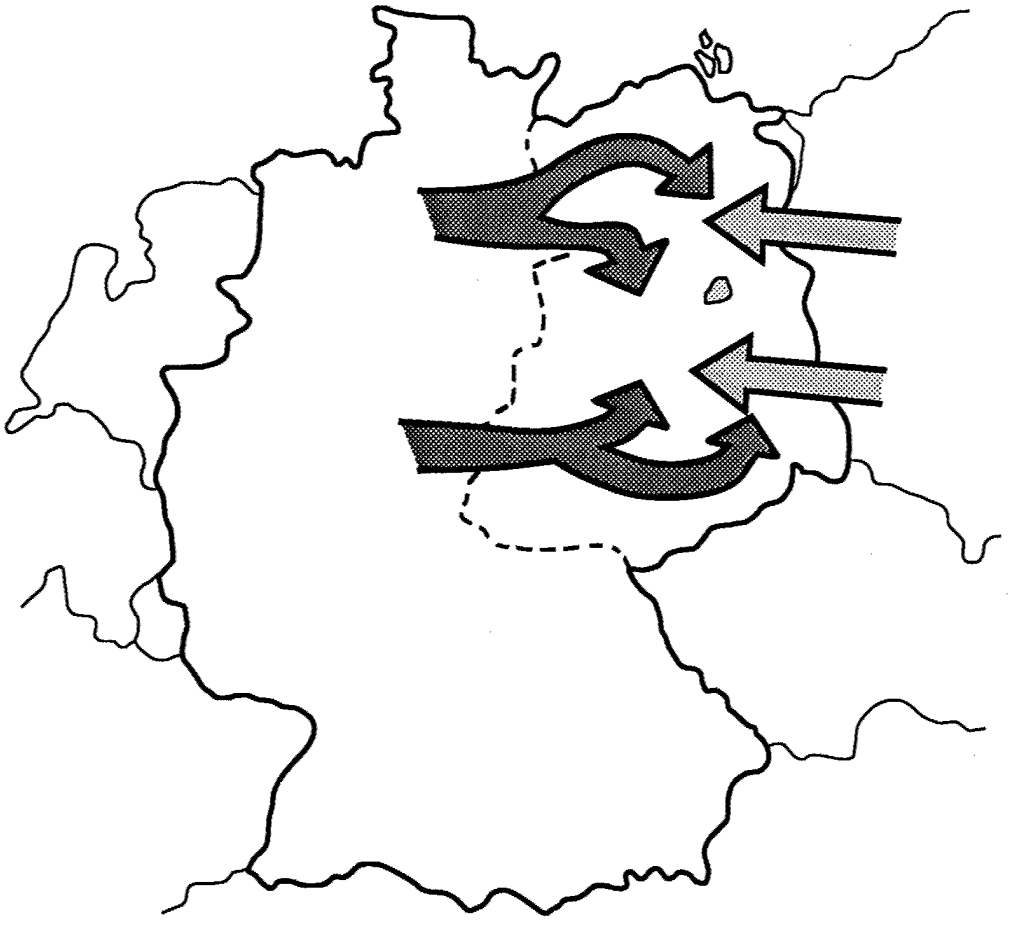

Fig. 26—Mobile Warfare Also Has Drawbacks

Second, this doctrine would compel NATO to design its forces and tactics for a form of warfare that can be interpreted as offensive in nature. Third, and equally important, this doctrine might not be militarily effective if NATO finds itself fighting outnumbered against an enemy that itself is well-schooled in maneuver and meeting engagements. Mobile defense is no military cure-all, especially for a coalition composed of several different nations faced with coordinating their operations on a fluid battlefield. It can offer a prescription for decisive victory, but it can also bring catastrophic defeat. Like linear defense and defense in depth, it brings problems of its own.

A Three-Tiered Echeloned Defense

NATO thus requires a doctrine that replaces linear defense with something less brittle, but one that itself is militarily effective and politically acceptable. A doctrine of an echeloned defense of eastern Germany, employing a combination of forward strength and mobile reserves, broadly accords with these requirements. NATO would shift its center of gravity into eastern Germany, thereby permitting a strong forward defense effort aimed at protecting the border regions. NATO's forces would not be deployed in a linear array, nor would they rely primarily on mobile warfare.

A frontal wall would be created along the Oder-Neisse rivers. Forces that otherwise would be deployed on less threatened terrain to the south would be redeployed to form large echelons of operational reserves. These reserves would operate in close conjunction with the frontally deployed forces. An echeloned defense with a frontal wall would deny the enemy an opportunity to open up the battlefield in ways that would permit the enemy's larger armored formations to overpower NATO's outnumbered forces. Meanwhile, NATO's forces would be massed together, creating a combination of forward strength and operational reserves that could block enemy thrusts. NATO's doctrine thus would be clearly defensive in nature, but it also would be both forward-oriented and militarily credible.

Two forms of echeloned defense will be examined here: a three-tiered and a two-tiered array. The manner in which a three-tiered echeloned defense would allocate NATO's forces on the battlefield is illustrated by Fig. 27. This doctrine would shift NATO's posture northward to protect the area that is most likely to be attacked. With only nine divisions now assigned to guard the less-threatened Bavarian region, fully four-fifths of NATO's forces would be available to defend eastern Germany. This change would provide enough divisions to form more than one brittle echelon.

With the exception of the area around Berlin, this employment doctrine establishes only a light screening force along Germany's border, even in eastern Germany. In total, 14 DEFs are committed to the immediate border areas, or only about 33 percent of NATO's totally mobilized posture. The remaining forces are withheld in the rear areas, where they would function as flexible operational reserves. This distribution contrasts markedly with a linear defense, which would commit 75 percent of NATO's forces to the border region and two-thirds of those units to the less threatened German-Czech border. It too enhances NATO's ability to form echelons in depth.

This doctrine would array NATO's ground forces in eastern Germany in what amounts to three separate tiers, one behind the other. The border region there would be divided into four corps sectors, averaging 75–80 kilometers. Apart from Berlin, which would be protected in strength, the Oder-Neisse rivers would be defended only lightly, with pickets of light infantry and reconnaissance forces. Backing up this frontal screen would be a powerful force

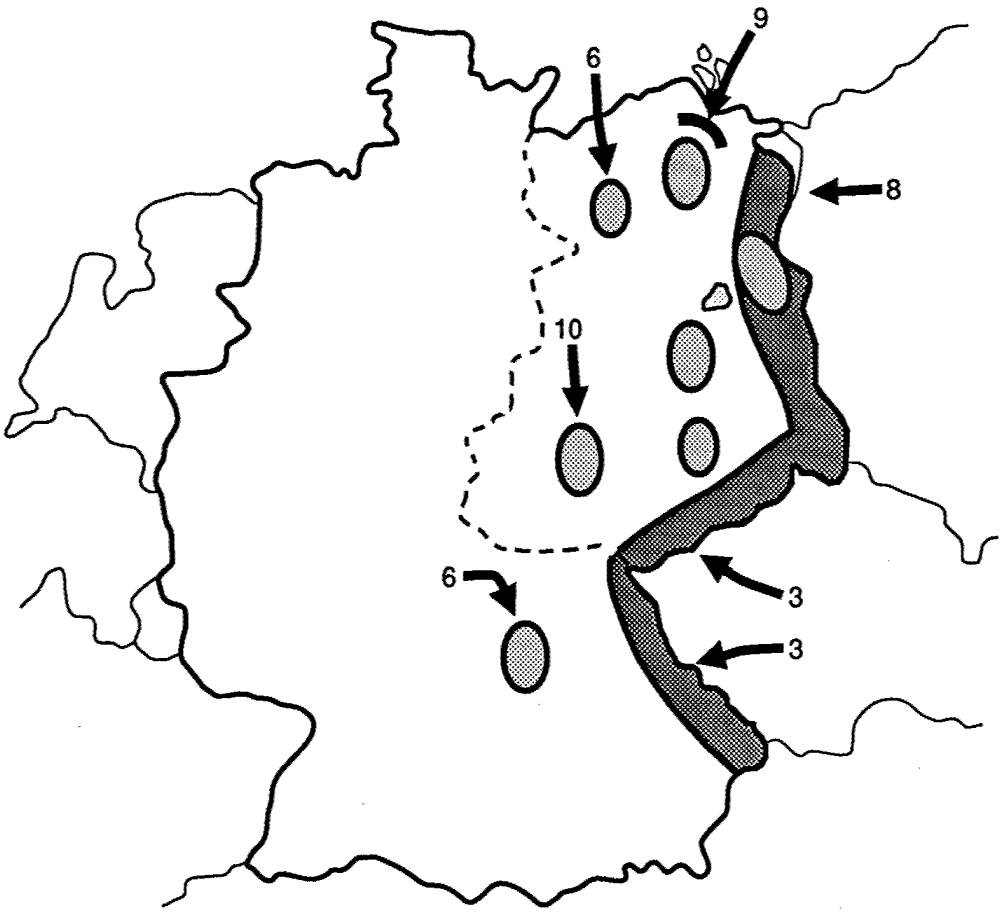

Fig. 27—Operational Concept for a Three-Tiered Echeloned Defense (DEFs)

of nine mobile armored/mechanized divisions (or helicopter assault units) in each corps sector. These divisions would be deployed well to the rear, some 100–150 kilometers from the Oder-Neisse river line. Further to the rear, at least 200–250 kilometers back, would be another echelon of mobile reserves roughly equal in strength to the forces deployed eastward. As envisioned here, these forces would be clustered into two full armies of six and ten divisions apiece. Alternatively, they might be organized into four corps-sized formations. Meanwhile, the equivalent of another large mobile corps would be deployed in Bavaria.

The central idea here is to construct a flexible employment doctrine that would respond to NATO's mobilization dynamics and be militarily effective. In a crisis, this force array could be assembled in phases. Especially in eastern Germany, the frontal screen would need to be established very early. Its mission during the mobilization process would be to shield the rear areas, where follow-on reinforcement activity would be conducted. Behind this shield, NATO's second echelon of mobile reserves would then assemble. Shortly therafter, the third echelon would assemble, thereby bringing the NATO's deployment operation to completion.

A pattern of sequential deployments would allow NATO to adjust to the time-phased availability of its forces brought about by different readiness levels and the later arrival of outside reinforcements. It also would facilitate traffic control, since NATO's forces could enter eastern Germany in a prearranged sequence rather than all at once. With demands for immediately deployed forces kept low, NATO's employment doctrine would not be vulnerable to coming unhinged because critical combat units were not available early. This change alone would mark a major improvement over the present array. Moreover, NATO's posture would still retain its flexibility and coherence if the Alliance failed to mobilize fast enough to establish a front wall along the border before an attack began. In that event, NATO's forces could begin combat operations in the area where the second echelon is deployed (or, for that matter, the third). As a result, NATO would be less likely to be caught off balance because of an ineffective mobilization effort.

Once fighting began, this doctrine would employ NATO's forces in the flexible manner that its employment scheme suggests. Above all, it would aim to cancel out the vulnerability to an enemy breakthrough created by an inflexible linear defense. Figure 28 displays how NATO might fight with this doctrine. NATO's three separate echelons would be employed in a mutually reinforcing fashion. The act of working together would help them to prevent a decisive enemy penetration, defeat the attack, and recapture any lost ground. The first two echelons of NATO's forces in eastern Germany would work together to conduct a nonlinear defense. The frontal screen would aim to prevent the Soviets from crossing the Oder-Neisse rivers quickly and easily. The second echelon of reserves could then move forward quickly and selectively to contain and destroy any bridgeheads across these rivers. If large Soviet forces did get across the rivers, NATO's second echelon forces would then give ground flexibly while destroying them with operational fires and counterthrusts against their flanks and rear areas.

NATO's third echelon would enter the battle if the forward forces became worn down or if advancing Soviet columns threatened NATO's rear areas. The forces in eastern Germany would focus on containing Soviet thrusts on either side of Berlin. They also would guard against a Soviet flanking advance through the Erzgebirge Mountain range. The forces in Bavaria, in turn, would protect against a Soviet drive there.

These forces would be flexible. They could operate directly in support of forward echelons by providing operational fires and maneuvers. They could swing north or south, adding further depth to threatened areas elsewhere. And once the enemy advance had been stabilized, they could counterattack the mission of destroying enemy pockets and restoring the German border. In performing their missions, these third echelon forces also would work closely together. For example, one cluster of these reserves might move into blocking position against a Soviet advance while another would launch a flanking attack against it. This kind of flexibility and capacity to work together would make NATO's posture an even tougher nut to crack.

A Two-Tiered Echeloned Defense

This doctrine also would concentrate NATO's forces in eastern Germany and deploy them in an echeloned fashion. But instead of establishing three echelons, it would form only two. It would bring forward the entire second echelon of the three-tiered defense and place it close to the border, creating a thick front wall, rather than a thin screen along the entire Oder-Neisse line. Meanwhile the third echelon would be brought forward to within 100–150

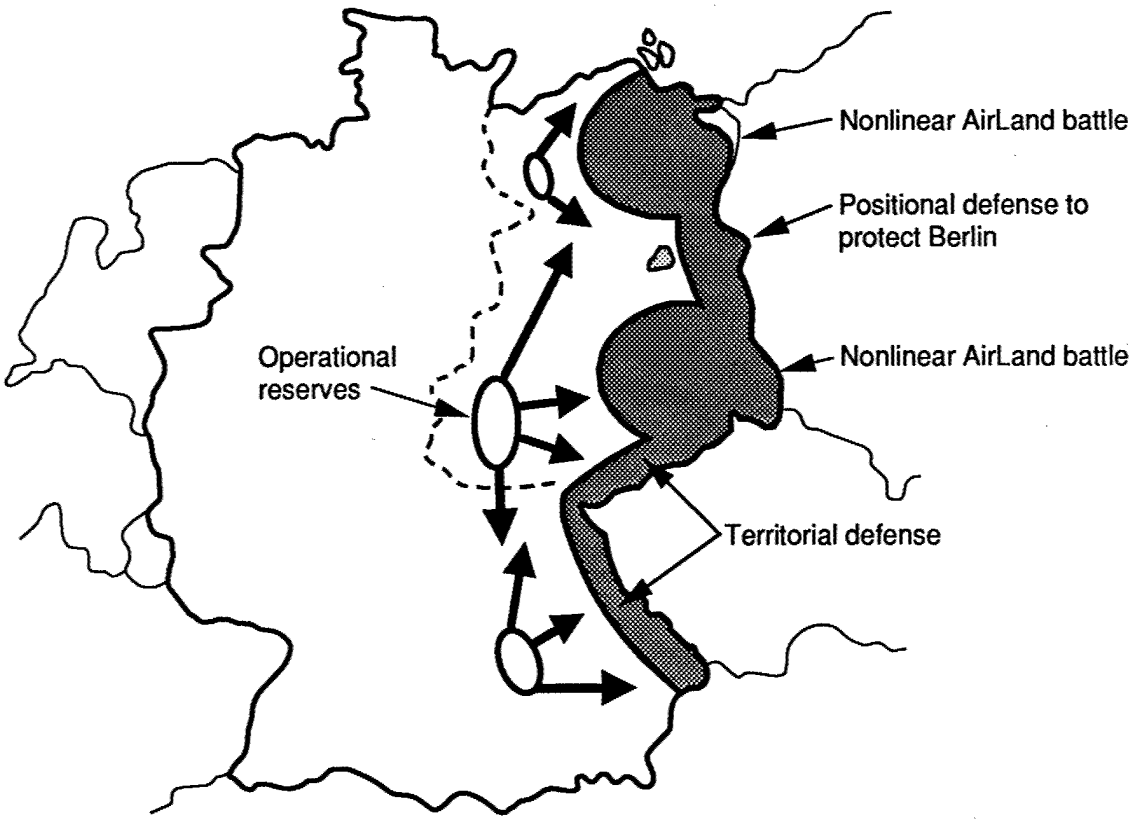

Fig. 28—Future Operational Concepts for Forward Defense

kilometers of the front wall, making a defense array of two echelons, with their center of gravity moved well forward. Figure 29 displays how this doctrine would deploy NATO's forces.

The goal of this doctrine is to perform the forward defense mission more effectively than its counterpart. The three-tiered doctrine might yield some German territory before stopping the Soviet advance. The two-tiered doctrine aims to deploy enough forces directly at the border to deny the Soviets any opportunity to make even limited territorial gains.

The two doctrines are similar in the sense that both rely on large echeloned reserves. Where they differ is in how they would conduct the initial stage of battle. The three-tiered doctrine would lure enemy forces onto German soil to destroy them by operational fires and flanking attacks. The two-tiered doctrine would place large forces in blocking positions, prepare the terrain, and aspire to destroy the enemy with shorter-range fires and local maneuvers.

This doctrine's forward defense effort could be conducted with a variety of tactical schemes. For example, all forces could be committed directly on the front line, or a strong-point defense could be erected that commits only one-half of the forces up front and holds the other half in corps reserve. In essence, the two-tiered doctrine would wage the first phase of war in a traditional way. Depending upon the extent to which enemy forces successfully

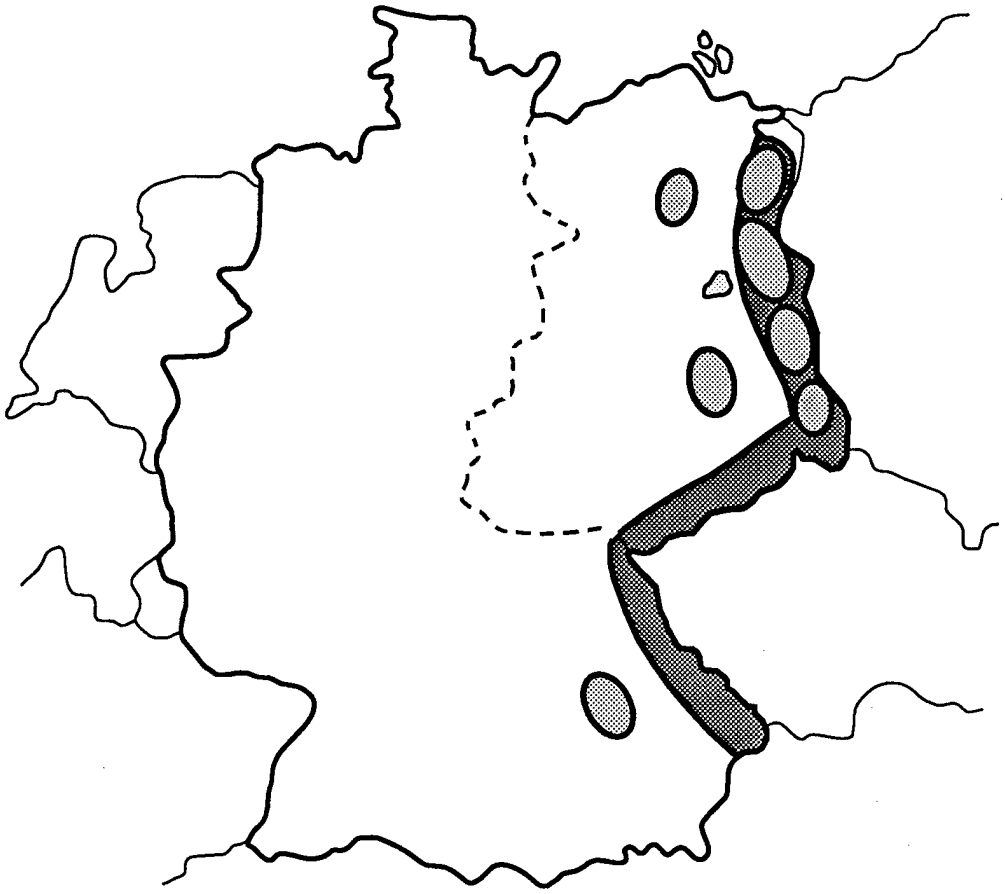

Fig. 29—Operational Concept for a Two-Tiered Echeloned Defense

crossed the Oder-Neisse rivers and advanced into German terrain, however, this defense easily could give way to a more fluid operation. These two doctrines might merge into one by being conducted in similar ways.

Compared with its counterpart, the two-tiered doctrine does hold out at least the promise of defending forward better. Its ability to deliver on this promise hinges on whether its tactics for protecting Germany's terrain are, in fact, more effective than nonlinear defense. The tradeoff here needs careful study. Whatever the case, this doctrine is more dependent on a successful NATO mobilization and reinforcement effort than its rival. Whereas the three-tiered doctrine is capable of reacting flexibly to a slow force buildup, this doctrine's rush to move forward could leave NATO's posture off-balance on D-Day. Its emphasis on building a strong front wall does provide greater scope for mechanized infantry forces, defensive defense, and other measures designed to strike a less offensive pose toward the east. NATO's choice between these two doctrines therefore would come down to a matter of priorities.

Regardless of which way the Alliance might tilt, both doctrines strive to replace NATO's brittle linear array with a more robust echeloned defense, which could help NATO attain the

goal of maintaining a stalwart conventional defense. Adoption of this doctrine also could help NATO achieve another important objective: making nuclear weapons an instrument of last resort. At a minimum, an echeloned defense would help eliminate the risk that NATO might be compelled to escalate very early because its frontal wall had unexpectedly been breached. The presence of sizable reserves as backstops would enable NATO to remedy this setback without having to call on nuclear weapons to shore up an otherwise unsalvageable situation. An echeloned defense would buy valuable time by delaying the decision to escalate.

Whether this doctrine would completely eliminate NATO's reliance on escalation would depend on more than its own robust qualities. A well-constructed echeloned defense can carry NATO only so far. It can solve glaring weaknesses that arise from maladroit plans and, to some extent, it can help patch over deficiencies in forces; but the posture that NATO does maintain in the post–Cold War era would still have to be strong enough to perform the missions that an echeloned defense demands.

Feasibility of Echeloned Defense

Will NATO's forces in the years ahead be sufficiently powerful to conduct an echeloned defense of eastern Germany? Assuming NATO's posture is reduced by about 10–25 percent, will the remaining forces be adequate for staving off an attack by 50–65 Soviet DEFs? Put another way, what are an echeloned defense's force requirements? And how do these requirements compare with a NATO posture that will be capable of building to 45 DEFs after full mobilization but could be as low as 30–35 DEFs on D-Day?

Rule of thumb planning factors would call for a NATO posture of about 40–45 DEFs. A force this size would be needed to cover the border regions in adequate density (with non-linear or frontal tactics) and to form a sizable rear echelon. Another useful yardstick is the force ratio. At lower force levels, NATO would need a 1:1 ratio to defend the disadvantageous terrain on the inter-German border. A somewhat different situation would apply in eastern Germany, provided NATO's forces there are large enough to generate sizable echeloned reserves.

NATO's defensive task there would be eased by the fact that the Oder-Neisse line is straight, fairly short, and marked by river lines that would slow an enemy advance. This advantage would allow NATO's reserves to counterdeploy more rapidly and effectively than along the inter-German border. NATO also would benefit from an advantage in the air since its tactical air forces are qualitatively better than the adversary's forces and because the Soviets would be compelled to operate from hastily prepared bases in Poland.

Especially since an echeloned NATO posture would not be vulnerable to crippling breakthroughs, it therefore should be capable of defending even if outnumbered by some margin. This study postulates that a 1.25–1.50:1 ratio in the enemy's favor is a reasonable standard for assessing NATO's force needs. Table 13 shows, this standard would call for a NATO posture of 33–52 DEFs, with a middle range of 40–43 DEFs.

To provide further insights, RAND conducted a series of dynamic simulations using the IDAHEX wargaming system. These simulations examined a NATO echeloned defense effort at alternative force levels on both sides and produced similar but somewhat more optimistic conclusions. The first simulation examined a defense effort conducted by German forces alone (16 DEFs). The German Army proved capable of containing an attack by 25 Soviet divisions, but because it was spread thin and lacked reserves, it was quickly overpowered by a 50-division assault.

Table 13

NATO Force Needs for a Manageable Force Ratio (DEFs)

	Required Ratio	
Size of Threat	1.50:1	1.25:1
65	43	52
50	33	40

The second simulation examined an echeloned defense effort conducted by 33–37 NATO DEFs against a 50-division Soviet assault. NATO's forces effectively contained this attack, but an echeloned defense proved critical to the effort. Confronted by strong NATO opposition on the Oder-Neisse rivers, the Soviet forces were incapable of advancing deeply into eastern Germany. A Soviet flanking operation through Czechosolvakia was blocked at the Erzgebirge Mountains by NATO's echeloned reserves. The battle was a close-run affair; timing, tempo and relative momentum proved critical. The combination of mobile reserves and frontal strength enabled NATO's forces to block enemy penetrations and to prevent the Soviets from gaining the initiative. The Soviet advance was stopped about 50 kilometers into eastern Germany and Berlin was saved. Subsequently, NATO was able to counterattack to restore lost territory. The battle was terminated at D+21 with the Oder-Neisse border restored and the force ratio stabilized at 1:1.

A less optimistic picture was portrayed by the third simulation, in which a larger Soviet threat of 65 divisions was modeled. This Soviet posture was able to overpower an echeloned NATO defense of only 33 DEFs and then advance deep into Germany. The combination of Soviet attacks across the Oder-Neisse rivers and through Czechoslovakia split NATO's defenses to an unmanageable degree. A NATO posture of 37 DEFs performed more effectively but not confidently. In some cases it was able to contain the attack; in other cases it buckled. A posture of 43 DEFs—enough to achieve a 1.5:1 ratio—proved necessary to consistently achieve a stalemate. Up to 48 total divisions were needed to restore the original borders.

Much additional analysis remains to be done on the issue of how many NATO forces are needed to defend eastern Germany. As interim measures, the three force-sizing techniques used here provide some useful insights. The NATO forces that are now planned to be available after 1995, these techniques suggest, can aspire to defend eastern Germany provided an echeloned defense doctrine is adopted. This sanguine conclusion depends on two critical assumptions. First, NATO's forces must not be allowed to fall well below the assumed level of 43–48 DEFs. Second, NATO must be able to mobilize and deploy these forces fast enough for all of them to participate in the battle. Violation of either assumption, especially both of them, would produce a more negative appraisal.

Although the issue of NATO's force levels is critical, so is the need to maintain an adequate mobilization and reinforcement rate. This requirement especially will be a necessary, if potentially neglected, feature of NATO's defense calculus if the Soviets prove capable of deploying a large, ready attacking force. The impending withdrawal of Soviet forces can give rise to the misleading impression that aggression against Germany would require such a

lengthy period of preparation that NATO can safely relax its mobilization guard. Some degree of relaxation, of course, will be possible. The Alliance would run the risk of taking this opportunity too far, however, if it ignores the post–Cold War factors that will be working against NATO.

These negative factors include NATO's own lower readiness and its greater dependence on outside reinforcement because fewer U.S. forces will be deployed in Central Europe. Furthermore, its ability to make prompt decisions will be inhibited by uncertainty and the longer distances that must be covered in moving into eastern Germany. The Soviets no longer will be able to mobilize and reinforce as rapidly as before, but neither will NATO. What matters here is relative capability. A Soviet attack on Germany launched after M+120 or later is one thing; an attack launched at M+60, or even M+90, is quite another.

Poland's status as a buffer state enters the military equation here, and in an important way. Polish resistance could delay any Soviet aggression appreciably; Polish acquiescence could speed the Soviets on their way. Likewise, NATO air operations in Poland (or even ground resistance) could buy valuable time for mobilizing and establishing NATO's defenses in Germany. In the years ahead, political considerations will very likely prohibit any alliance with Poland or a NATO decision to develop its forces for the express purpose of operating there. Nevertheless, the incentives for developing NATO contingency plans to enter Poland in wartime, at the Polish government's request, are strong.

Key Sensitivities

The estimates provided here clearly are tentative and will require further in-depth study. The preferred NATO employment doctrine might vary as a function of preparation time available. A short period might prevent NATO's forces from reaching Germany's eastern borders, thereby compelling a form of mobile defense. A longer period (about 60 days) would permit the kind of echeloned defense envisioned here. A very long period (e.g., 150 days) might permit the kind of entrenchment efforts that could make a more linear array quite hard to breach. To some degree, employment doctrine is sensitive to reaction time.

NATO's required force levels are also sensitive to external events. For example, a long buildup period allowing for major entrenchment efforts could lower NATO's overall need for ground and air forces. Similarly, major NATO air operations in Poland before the Oder-Neisse battle might inflict enough damage on advancing Soviet forces to reduce NATO's total force requirement. At a minimum, it might shift NATO's preferred posture toward a larger mix of air forces and fewer ground forces. Conversely, a larger Soviet threat than estimated here could drive NATO's requirement upward.

FUTURE U.S. AND ALLIED MILITARY CONTRIBUTIONS

A NATO decision to adopt an echeloned doctrine for defending eastern Germany would carry with it a need to adopt a new and effective scheme for allocating missions to member nations. For all the reasons mentioned earlier, recreation of the layer cake would not make sense. But what new coalition arrangement should take its place? And what are the implications for future U.S. and Allied contributions to NATO's defenses? This study examines two options, each of which responds to its own sense of NATO's priorities.

Multinational Forces

The first option, displayed in Fig. 30, might be labeled a "multinational" approach to force planning. It has two prominent features. First, forces from at least three nations would be committed to the forward zone (with a nonlinear doctrine, their units would deployed somewhat further to the rear). NATO's front wall on the Oder-Neisse rivers would be made up of two German corps, a U.S. corps, and a British corps. Each of these national corps might have a multinational dimension by having one brigade or division from another nation assigned to it. For example, one of the German corps might have a U.S. division, and the U.S. corps might have a German division. Alternatively, the British corps could be a fully multinational formation with one British, one Belgian, and one Dutch division. Second, NATO's rear echelon also would take on a multinational character. All three clusters of reserve forces would be composed of forces drawn from anywhere between two and four nations, with multinational corps composed of two divisions from one nation and a third division from another.

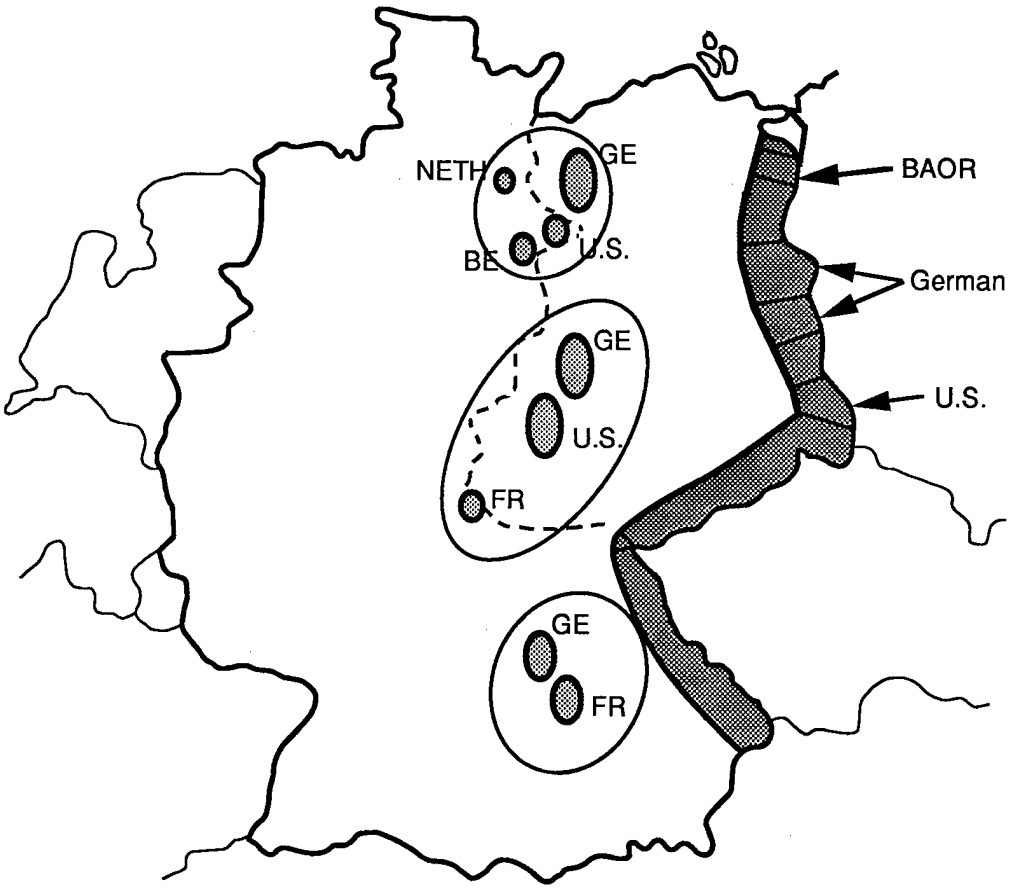

Fig. 30—A Multinational Approach for Building an Echeloned Defense

This scheme has two main goals, and it promises to do a good job of pursuing both of them. First, it takes aim at cementing NATO's unity in the post–Cold War era. In particular, this option will reassure Germany by securing the kind of military commitment from its partners that signals continued NATO constancy. By assigning frontline roles to several nations other than Germany, it calls on them to commit themselves unequivocally to defending eastern Germany. In doing so, it avoids any hint of leaving Germany uncertain about NATO's support for this mission. This reassurance is intended to help dampen any sentiment within Germany for disengaging from the Alliance because NATO could not be relied upon in a moment of danger.

Additionally, this option sidesteps any appearance of singling Germany out. It assigns Germany a combination of frontline and rear area roles, as it does other NATO nations. Germany and all its NATO allies would continue to assume obligations and burdens that are alike. Consequently, Germany's forces and strategy could not readily be isolated for special treatment by outside nations in arms control negotiations. Nor could Germany easily be discriminated against in other ways.

Beyond reassuring Germany in a fashion underscoring that nation's status as a co-equal member of the Alliance, this option would tie Germany's forces to NATO's integrated military command in ways that inhibit them from operating on their own. This feature would help reassure the Allies about Germany's own intentions in a crisis. It also would impress upon all members that they continue to play important military roles in NATO, roles requiring them to continue meeting their military commitments to the Alliance. Finally, it assigns prominent frontal roles to American and British forces, thereby enhancing the credibility of the extended nuclear deterrence that these two nations provide. This option would help perform many of the layer cake's important political functions, thereby helping keep the transatlantic bargain alive.

Second, this approach to organizing NATO's posture prods the Alliance into taking the next step toward integration by forming truly "multinational" ground combat formations. This concept picks up on an idea that was once taken seriously and then was cast aside and ignored for over 30 years. Now that the political climate has changed, it has been resurrected in an updated form. The concept of multinational formations was looked at with favor in the early 1950s because it offered an acceptable way to rearm West Germany and to bring it into NATO. It now has reappeared as a device to help keep Germany in NATO, to provide a way to maintain NATO's forces on German soil, and to discourage NATO's members from reducing their forces too far. Its rationale thus is heavily political.

This idea has not come back to life entirely for pragmatic reasons. The military arguments favoring this idea are powerful and have been recognized in NATO circles for many years. NATO's defense posture has long suffered militarily from its failure to pursue coalition planning in more ambitious ways. The historical fact that NATO has not yet traveled very far in this direction is attributable more to a lack of willpower than to any policy judgment against it. The result has been redundancy, duplication of effort, incompatible weapons, unharmonious doctrines, overly competitive armaments programs, and a host of other deficiencies. The idea of creating multinational units carries with it the hope that some of these problems might be solved.

Quite apart from these military incentives, larger policy considerations argue in favor of this course. The creation of multinational formations could help propel the West European nations further down the road toward a common defense establishment and military self-sufficiency. Although some skeptics might question the wisdom of pursuing this goal, the

continued disputes between the United States and Western Europe over burden-sharing argue in its favor. In any event, West European military integration eventually will occur as a byproduct of efforts to achieve economic and political unity there. Multinational formations encourage this evolution to take place under NATO's umbrella rather than within the European Community, the WEU, or some other organization. They thus offer a vehicle for joining the West Europeans together in ways that keep the United States militarily involved in Europe and the transatlantic alliance intact.

In theory, multinational integration can be pursued at a variety of levels and with different degrees of ambition. At one extreme, it can be implemented in name only: for example, by attaching the NATO label to an all-German corps. At the other extreme, it can be implemented at the level of the individual soldier: Infantrymen from several nations can be brought together to form a squad. As envisioned here, it would be pursued only down to the division level. That is, individual divisions would retain their national identity, but they would be organized into multinational corps formations that would operate under NATO's flag, take orders from a NATO commander, and have a common doctrine and logistic support system. These corps formations would be multinational in more than name only.

This option would take more than a symbolic bow in the direction of multinationalism. Particularly affected would be the British, Dutch, and Belgian armies, which might be reduced to the point where they no longer are large enough to form full corps. The German, American, and French forces would be less involved, but neither would they be left out. All told, about 50–75 percent of NATO's posture would become multinational. Along with this change in combat forces could come a major alteration of NATO's posture in other areas. NATO's command structure and logistic system might become more multinational, as could NATO's training and exercise programs. Eventually parallel adjustments might take place in NATO's policies for promoting standardization, armaments cooperation, and common industrial ventures. All these changes would propel NATO not toward the dissolution that many expect since the Cold War has ended, but in precisely the opposite direction.

The multinational defense option thus has attractive political strengths, but it also carries with it two important military drawbacks. First, it might weaken NATO's ability to mobilize and reinforce rapidly in a crisis. Because it assigns frontline roles to several nations, it would place pressure on NATO to mobilize the Alliance's entire posture early in a crisis, rather than German forces alone. This feature could increase the likelihood that NATO would become mired in a political debate about the wisdom of taking this major step, thereby delaying its mobilization decision. Even if NATO did make this decision promptly, moreover, the physical act of building a frontal echelon might be more time-consuming than if German forces alone were entrusted with this mission. This especially would the case if U.S. and British forces based on the continent were not large enough to form corps-sized defenses and required outside reinforcement. Dutch and Belgian forces, which would have to travel a long distance from their homelands, also might have trouble reaching the forward areas quickly. The delays brought about by trying to move these forces might not matter in a drawn-out crisis that provided NATO ample time, but a quick-breaking crisis might be another matter.

This option might weaken NATO's ability to fight proficiently. It would divide German, U.S., British, and French forces into segments, deploy them at widely separated locations, and require them to perform dissimilar missions, which would undermine the ability of these national armies to achieve the mass, concentration, unity of command, and common doctrine that are critical to fighting well. This particularly would be the case in situations where

these national armies could not, at a minimum, operate in cohesive corps formations. Since only about 25–50 percent of NATO's posture would be organized into national corps structures, the Alliance would be left dependent on the capacity of its multinational formations to perform at a high level.

Unfortunately, multinational units are themselves a doubtful proposition because of their complex composition. Differences in language, doctrine, tactics, weapons, and support systems easily could undermine their combat effectiveness. These barriers, in theory, can be surmounted through training, exercises, and other hard work. The period of more relaxed European political relations ahead might provide NATO breathing space to make progress in this area, but NATO's history hardly provides grounds for confidence on this score. Indeed, previous attempts to launch ambitious multinational programs normally failed, largely because national sovereignty got in the way. The successes that were achieved were minor. If NATO were to fail to breathe military life into its multinational formations, half of its posture might wage war less effectively than surface appearances suggest.[10]

A critical issue is whether the losses in combat power suffered by pursuing multinationality would be marginal or major. Much would depend upon the extent of multinationalization and the degree to which compensatory training and standardization measures were adopted. Without effective offsetting measures, however, NATO's units might experience as much as a 20 percent dropoff in combat effectiveness. A loss of this magnitude could have potentially serious negative consequences for NATO's conventional defense strategy. For example, a multinational NATO posture of 45 DEFs might appear adequate on paper but actually be capable of fighting with the strength of only 36–40 divisions. As the wargaming analysis presented earlier suggests, this could transform a highly confident posture into one of only medium confidence. The consequences could be even more severe with a smaller posture; in that case, multinationality plausibly could spell the difference between success and failure. These risks would need to be taken into account in evaluating multinationality and the defense programs that are to accompany it. Militarily, NATO should pursue multinationality only if it is prepared to implement the programs needed to ensure that the exercise does not produce a hollow defense posture.

Specialization

The risk that NATO's buildup might be slow and that its forces might not fight well calls into question the military wisdom of this option, its political attractions notwithstanding. The second option to be examined here endeavors to solve these military problems by pushing NATO in the opposite direction of specialization. It does not ignore the idea of multinational integration; it provides some room for movement in that direction. But it does not make this idea the centerpiece of NATO's defense plans. The manner in which it would deploy NATO's forces on the battlefield is displayed in Fig. 31.

This option would build an echeloned defense on a foundation composed of strong national armies. It assigns the task of erecting NATO's frontal echelon(s) primarily to the German Army. It gives the other Allied armies the function of forming NATO's backup echelon and operating as mobile reserves. The U.S. Army would play a particularly important role by occupying the center position in NATO's reserve layout. On the left would be a force dominated by the British Army, but with a strong multinational flavor. On the right would be the French Army, which would operate from Bavaria. The net effect would be to

[10]For a historical discussion of NATO's efforts to create multinational forces in the 1950s, see Karnowski, 1984.

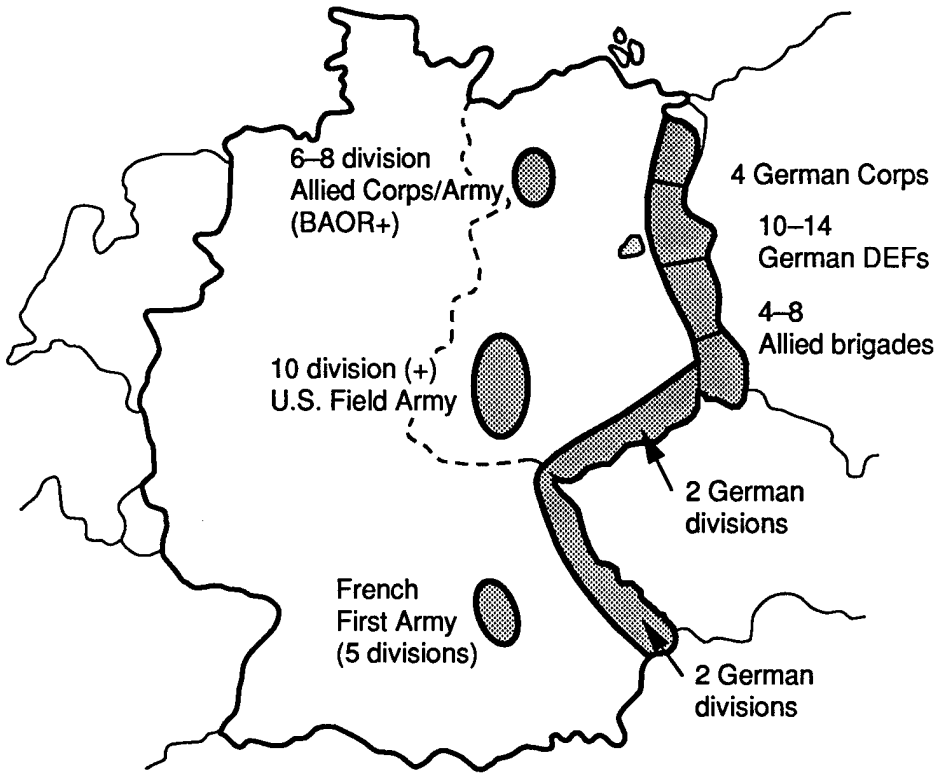

Fig. 31—Specialization as an Approach to Building an Echeloned Defense

establish an echeloned array dominated by identifiable national formations rather than homogenized multinational units.

One important goal of this concept is to improve NATO's capacity to mobilize and reinforce in a rapid, orderly manner. This concept orients the German Army to the frontal defense role because, with some units already based in eastern Germany and the others readily available in western Germany, it will be best situated for reaching the forward areas quickly. The act of first sending German troops in unison also might ease what otherwise could be a difficult traffic control problem. Also, if NATO were to find itself mired in debate about whether to mobilize fully, the German Army and Air Force could be detached and moved forward separately. Other NATO forces could follow later, once a decision to mobilize fully had been made and these forces were ready to move forward. With a front wall thereby built early, NATO could debate the larger issues in the confidence that further delay would not bring about its undoing.

This concept has a second goal of building military strength through functional specialization. Each of NATO's national armies would perform a limited set of functions. It would achieve operational coherence by combining these functions in a way that produces a synthesized plan for conducting a robust echeloned defense.

Each of NATO's national armies would be able to mass in one location and to operate under a single command, with a common doctrine, on behalf of a clearly identified mission.

The German Army would focus on defending its nation's border areas, makes political sense. The German Army already is well-endowed for this mission, and it would now be able to fine-tune itself further. Conceivably it could adopt some defensive defense concepts (e.g., fewer tanks and more mobile infantry), which would have a salutary political effect across Europe. Meanwhile NATO's other national armies would be left free to tailor themselves for the operational reserve mission, which makes political and military sense for them. In particular, the U.S. Army would now be largely disengaged from frontal defense tasks and could specialize in giving NATO the capability for mobile warfare that it is best endowed, by doctrine and weapons, to provide. The net effect on NATO could be highly positive if each nation army, in fact, performed its assigned function. With a posture anchored on German forward strength and American reserve power, supported by other national forces playing a contributing role, NATO would stand a better chance of rebuffing aggression of any sort.

The principal drawbacks of this approach are that Germany might be uncertain about NATO's constancy and sacrifice the political advantages of a multinational posture. This option consequently provides for a limited degree of multinational integration, especially in highly visible areas. In particular, some four to eight Allied brigades (including U.S. forces) would be assigned to the German frontal defense and scattered across Germany's eastern border. Commitment of these forces would help assuage German concern about the Allies' reliability in a crisis. A few German brigades would be withheld in rear area roles, thereby reducing the risk of singularizing Germany, but still providing considerable links between German forces and other national armies.

Most of NATO's large combat formations would acquire a multinational dimension at least in limited ways and could plausibly be labelled "multinational" units. This step, of course, would be undertaken for appearance's sake and would not alter what would still be a coalition of national armies. But provided that it did not go too far toward stressing form at the expense of content or downplaying national sovereignty, it might add useful political symbology to NATO's military arsenal.

NATO would still be able to form one or two truly multinational corps by bringing together the small British, Dutch, Belgian, and other national units that would be deployed in the northern rear areas. Indeed, NATO would be militarily compelled to take this step to forge these units into a coherent whole. Some American units could join this enterprise, along with the French if possible. The net effect would be that about 20 percent of NATO's posture would be in explicitly multinational formations, a greater percentage than exists today. Also, NATO would have ample scope for pursuing the idea of integrated multinational logistics forces, which will make sense regardless of whether NATO's combat forces specialize.

The entire military posture would remain under NATO's integrated military command structure. The task of designing an echeloned defense of eastern Germany probably will compel a major change in the present command system. The division of AFCENT into two geographically oriented Army Groups—NORTHAG and CENTAG—will make less sense in the years ahead. One option for replacing it would be to put NATO's various corps formations directly under AFCENT's command. If this option creates an unmanageably large span-of-control problem, one or two subordinate separate Army Groups could be created: a "forward" command and a "reserve" command. Alternatively, four separate "armies" could be created: a large one forward and three small ones back, each identified by number rather than by geographic or functional affiliation. The main difference here is that whereas an "Army Group" typically is composed of 4–5 corps and 12–15 divisions, an "Army" normally

has three corps and 9–10 divisions. These options need to be studied carefully, with a view toward balancing the complex requirements involved. Regardless of which is chosen, NATO would remain an active military coalition, with a strong multinational flavor, by virtue of its integrated command structure alone.

Whether 20 or 70 percent multinational integration would be sufficient to meet NATO's political requirements would have to be debated. In any case, NATO will face complex tradeoffs in addressing how its defense posture should be pieced together. Many of NATO's political goals bearing upon internal unity pull in the direction of multinational initiatives, but military exigencies push toward specialization. Broader political trends in Europe seem likely to cut both ways. Western Europe's drive toward economic and political integration creates incentives for national defense establishments to work closely together. Nevertheless, Britain, France, and the Low Countries will not be directly threatened by the Soviet Army and therefore might prove less willing to assume frontal missions and entangling commitments than before. How these competing trends will play out remains to be seen.

In all likelihood, NATO will serve its cause best by pursuing a combination of measures, some that integrate and others that specialize. Any single-minded pursuit of multinational integration would run up against powerful military barriers and adverse political trends. NATO can hardly afford to totally embrace a concept that might leave its defense posture fragile, nor can it wholeheartedly call for greater resolve at a time when tensions in Europe are cooling. But even in the interest of military efficiency, NATO cannot afford to specialize in ways that could undermine its political cohesion. For these reasons, a balanced mix of integration and specialization makes sense.

A balanced plan for the future would call on NATO to pursue multinational forces in a modest way that first tests the waters. For example, several multinational corps could be formed with a safety valve that permits transfer back to purely national corps in a crisis. Creation of a single, highly integrated testbed multinational corps, one not at the heart of NATO's warfighting plans, might be a good place to start. With such a limited effort underway, NATO might find itself with greater latitude for pursuing the path of specialization. These two initiatives might complement each other, the one satisfying NATO's political needs and the other its military requirements. As time passed by, NATO could adjust its balance by creating more multinational units, or by specializing more, or by pursuing some combination of the two.

Future U.S. Military Role in NATO's Conventional Defense Strategy

NATO has long had a quite different plan for defending the Northern and Southern flanks than the Central Region. NATO's plan for the flanks has called on member nations there to defend themselves initially until help could arrive. Although this plan placed a heavier burden on frontline states than in the Central Region, it did serve its purpose. Military requirements alone will prevent NATO from applying this model to Central Europe in a wholesale way. Nonetheless, political trends—including Germany's preference to defend itself coupled with the desire of other West European nations to relax their defense efforts—might call for some movement in this direction.

NATO will still need to engage in combined planning to maintain a smaller but adequately large, reconstitutable defense posture. It will achieve this goal only if nations on both sides of the Atlantic continue to contribute their fair share of forces. Clearly the current

German Army alone will not be able to defend its borders against a major threat in the absence of sizable military help from other NATO nations. Asssuming they are not reduced beyond present plans, the forces committed by other West European nations could help remedy the situation by increasing NATO's total posture to about 25–30 DEs, but this posture would still fall well short of adequacy. For this reason, a sizable U.S. military contribution, one provided primarily through outside reinforcement, will still be necessary.

This reliance on U.S. nuclear and conventional military strength is not likely to change even though Western Europe will make further strides toward political integration in the coming decade. As matters now stand, the West Europeans have no plans for, or official interest in, creating a continental security pact that could replace NATO. The European Community in late 1990 endorsed the idea of creating a "West European security pillar," but not one that will replace NATO. Nor do these nations seem inclined to use their growing economic strength to expand their forces by any appreciable margin. If any change comes about, it is likely to be in the opposite direction. Some nations—Belgium, the Netherlands, and Denmark come to mind—may even choose to pare back their forces beyond the point envisioned here or even loosen their military ties to NATO. These measures would place pressure on the German government to expand its own military forces to make up the difference. But even if Germany acted accordingly, the original equilibrium would merely be restored.

The most likely scenario is that West European forces will decline somewhat in future years (by 10–25 percent) and then stabilize at a level that is acceptable in political and budgetary terms. These forces will need to meet the readiness, modernization, and sustainability standards that NATO's official force goals set for them. If they meet these standards, they will continue to provide the backbone of NATO's readily available posture, but they will not be strong enough to support NATO's military strategy on their own.

Germany and the NATO Alliance will therefore continue to need the United States. What, then, should be the U.S. military contribution to NATO and to European stability? The answer to this important question currently is unclear, and much will depend on how a host of changes at work in Europe play themselves out. Although political requirements will figure importantly in the calculus, so also will military strategy issues. Both now and in the future, these issues merit close scrutiny, particularly in Central Europe. The United States will continue to face military requirements on the northern and southern flanks, needs that might grow if the Middle East continues on its troubled course. But Central Europe, and the specific defense requirements that it creates, will remain the heart of the matter.[11]

Two issues are involved here. First, how many U.S. forces should be kept in Central Europe? Second, how many reinforcements should the United States plan to provide? Of the two, the former question is likely to attract the most public attention, but in terms of its effect on the U.S. defense posture and budget, and potentially Europe's security, the latter question is important too. Both questions will need to be answered before the future U.S. role in Europe can be defined.

It is easy to forget that Soviet forces will remain in eastern Germany for several years. During this period, the United States will need to retain a counterbalancing military presence in western Germany. After Soviet forces have fully withdrawn, a different calculus will emerge, one based on larger strategic requirements for fostering overall stability in Europe. But there will still be a powerful rationale for maintaining a sizable peacetime presence there.

[11]For an official assessment of the future U.S. defense budget and force structure, see Cheney, 1990. For an alternative view, see Kaufmann, 1990.

Perhaps the simplest, but most potent, political reason for keeping some U.S. forces in Europe is to provide a visible reminder to all nations that the United States will remain a European power with vital interests there. Earlier this century, two world wars broke out in Europe partly because this elementary fact was overlooked by Europeans and Americans alike. The Cold War, in turn, was won because all sides came to regard the United States as a potent actor in European security affairs. A continued U.S. presence would help ensure that these lessons of history are not forgotten.

Another political reason for maintaining a U.S. presence is to reassure Germany and other allied nations regarding American constancy, thereby to help maintain NATO's unity and a sound transatlantic relationship. A continued American military presence would also send a strong but unprovocative message to the Soviet Union about the need for responsible conduct on that nation's part. In other words, as long as stability in Europe remains a goal to be achieved through the preservation of a balance of power, a U.S. peacetime presence will be a necessary ingredient for maintaining a sound security architecture there.

Although the risk of war will be low in the years ahead, keeping U.S. theater nuclear forces, especially air units, in Europe will help maintain a NATO nuclear deterrent. Because they show resolve, U.S. ground forces also play a role in the extended deterrence equation. But with the U.S. conventional role likely to shift toward providing reserves, American forces no longer will perform the major frontline missions that formerly demanded the presence of many combat and support units. This change will permit a sizable U.S. drawdown.

A total withdrawal will not be possible because U.S. forces are likely to retain at least residual obligations in NATO's defense plans that will demand the continued deployment of some units. A peacetime presence will provide an early U.S. response to possible crises and serve as a vanguard for a later reinforcement effort. Some U.S. forces will need to remain in Europe for joint exercises with Allied forces and the development of common doctrine. Further, a visible U.S. presence will be an important factor in helping to preserve an appropriate role for the United States in NATO's command structure. Although American military requirements in Europe will be less than during the Cold War, they will not disappear entirely.

The question of how many U.S. forces should stay in Europe after the Soviets have withdrawn should be addressed in the framework of these requirements. In Central Europe, U.S. forces should remain large enough to provide at least a credible increment of combat power: e.g., an Army corps and a few tactical air wings. They should also provide the support infrastructure to permit prompt outside reinforcement in a crisis. Growing tensions in the Mediterranean and the Middle East could demand a modest increase in U.S. force levels on the southern flank.

For many years, the United States has maintained a similarly functioning presence in South Korea with a manpower level of only about 40,000 troops. NATO and Europe create a more complicated military environment, however, so the U.S. presence will need to be higher than in Korea. For instance, the U.S. presence in Korea has normally amounted to one division and two air wings. A combat posture in Central Europe of double this size, coupled with deployment of some forces on the southern flank, presumably would drive the manpower requirement up to about 100,000. The need to retain specialized command and control, intelligence and communications, logistic support, infrastructure maintenance, and other units could drive this requirement higher yet. In the end, U.S. requirements probably will be substantially lower than the 225,000 figure (195,000 in Central Europe) that has been contemplated in official plans. The exact figure will need to be determined through careful, detailed analysis.

The question of U.S. reinforcement levels should also be subjected to careful analysis, particularly whether U.S. Army equipment at POMCUS sites should be withdrawn as American combat forces depart. This issue turns on whether U.S. Army and Air Force reinforcements would need to return in a rush should a crisis occur. At first glance, the upcoming shift of any military D-Day outward to two or three months after Soviet mobilization seems to provide greater scope for relying on sealift, rather than POMCUS, for deploying U.S. forces to Europe. As discussed earlier, however, NATO plausibly could still find itself facing a reinforcement race were it to appreciably delay its own decision to mobilize. Retention of a few POMCUS sets in Europe (e.g., 2–3) would provide a useful, low-cost hedge against this risk. As long as NATO remains willing to provide infrastructure funds to pay for POMCUS maintenance, their complete return to the United States would become necessary only if the Army found itself short of equipment.

Careful analysis also will be needed in determining exactly how many U.S. ground and air reinforcements should be sent to Europe in a future crisis. In years past, the Department of Defense planned to send the equivalent of 18–20 Army divisions (active and reserve component) and 22 USAF wings as reinforcements from the United States. In the years ahead, the requirement for U.S. forces will dip downward as the defense situation in Europe improves. Nonetheless, the analysis presented here suggests that, in addition to the U.S. forces remaining in Europe, some 10–15 divisions and 12–17 air wings will be needed to defend eastern Germany. This requirement would rise or fall with the Soviet threat. Thus, the reinforcement mission will continue to be a demanding one.

The challenge ahead will be to configure the U.S. military posture to match NATO's new reinforcement timelines and to play the role of a mobile reserve in defending eastern Germany. For the U.S. Army, these requirements call for an emphasis on nonlinear warfare, modern equipment, and agile combat formations. To the extent that global missions will prevent these requirements from being met by active units, greater emphasis will need to be given to using reserve component units. Although U.S. forces will be reduced in the years ahead if the defense budget declines, the posture that is planned for the late 1990s seems adequate to the task. Further cuts could bring the reinforcement mission into jeopardy. How the United States reacts will depend on its budgetary and strategic priorities in the mid-1990s and beyond. Unless stability in Europe becomes a condition of nature rather than a goal to be achieved, the quest for security there seems likely to remain a weighty factor in U.S. policy.

4. LINEAR DEFENSE IN HISTORICAL PERSPECTIVE

The main argument for devising a new, more flexible theater employment doctrine, culminating in an echeloned defense of the Oder-Neisse line, is that a linear defense at much lower force levels than now would leave NATO vulnerable. A Soviet attacker could concentrate along NATO's front line, gain a major breakthrough because NATO would lack sufficient reserves, and then prosecute a maneuver war in NATO's rear areas, possibly resulting in a decisive Soviet victory. This risk would be especially high if NATO finds itself still outnumbered by some margin as expected, but it would still be serious even if the future Central European force balance turns out to be a more favorable condition of parity.

This section is a historical inquiry surveying the up-again, down-again fortunes of linear defense at parity in the modern era. Section 5 is a formal theoretical and quantitative inquiry of how a linear defense would function against a concentrated enemy attack and how NATO's declining force levels in the years ahead will erode its capacity to continue relying on a linear doctrine. These two sections attack the issue from different angles of vision, and the conclusions they reach are mutually reinforcing: They provide additional analyses on the military reasons why linear defense can become vulnerable if NATO's forces are not well prepared.

The purpose of the historical review is not only to show that a linear defense often has been beaten in the past because the defender was unable to execute his own doctrine, it is also to shed analytical light on the determining factors that caused linear defense to fail in some cases but to succeed in others. The goal is to illuminate the conditions under which linear defense can become a source of weakness rather than strength, and vice versa.

Eight historical cases in which a parity fight took place between a breakthrough-seeking attacker and a linearly arrayed defender are drawn from World Wars I and II, the last two times a theaterwide campaign was fought on the Central European terrain. They provide clear examples of how battlefield outcomes can be shaped by the interaction between the offensive and defensive strategies being studied here.

These battles do not purport to offer a representative sample from which generalizations can be drawn regarding the statistical distribution of outcomes along the spectrum of victory and defeat. They do offer a range of outcomes that covers both extremes and the middle. In five cases drawn from World War I, the defender strongly rebuffed the attacker. In two cases drawn from World War II, the attacker routed the defender. In the final case, the attacker initially made progress, but the defender rallied to save the day. These cases, shown below, provide a good database for empirically analyzing the causal factors at work, and their interplay, in breakthrough battles.

Defender Success	Mixed Outcome	Attacker Success
Ypres I and II	Battle of	1940 Battle of France
The Somme	the Bulge	Barbarossa
Paschendaele		
German offensive of 1918		

LEGACY OF THE 19TH CENTURY

The concept of an offensive manuever strategy is most commonly associated with the German Army's "blitzkrieg" doctrine of World War II, but the idea had its origins, and was applied in various guises, long before that. It first appeared as a technique of modern warfare during the Napoleonic era. From then on, it played a central role in the military doctrines of many European and American armies. The Wehrmacht updated the idea to meet the conditions of the mid-20th century.

Napoleon himself was not the creator of this strategy. Attacking columns crashing through and encircling defending lines goes back a long way. Over the course of military history, renowned military commanders rode this strategy to fame, among them Alexander, Marlborough, and Frederick the Great. Napoleon brought this strategy into the modern era by harnessing the power of the French nation-state to support it. By doing so, he was able to conquer nearly all of Europe, defeating a long string of opponents over ten years.

Napoleon developed the modern army: a large, fast-moving, offensively oriented machine that could strike quickly and with devastating effect. Motivated by nationalistic fervor and led by skilled generals, Napoleon's *Grande Armée* was composed of multiple corps formations, all operating independently but under central control on behalf of a common offensive objective. Napoleon mastered the art of blending together artillery, infantry, and cavalry, thereby enabling the French Army to attack in a coordinated sequence. This capability, new to Europe, helped Napoleon score a long list of decisive victories throughout his career, normally against equally large opponents.[1]

Napoleon's operational concept explicitly rejected positional warfare, set-piece frontal attacks, and lengthy attrition battles. He aimed to destroy the opposing army quickly, thereby breaking the enemy nation's psychological will to resist. The bold offensive stroke, based on speedy maneuvers aimed at deceiving the enemy, became his principal means to this end. Normally Napoleon sought to draw his opponents into battle from the march, with both armies moving forward to engage one another. Typically French forces approached the battlefield deployed across a broad front, thereby luring the opponent into dispersing his own forces along an extended line. Some French forces would then feint against the enemy's main elements in an effort to lure them into battle and pin them down. The bulk of the French Army meanwhile would rapidly concentrate elsewhere and attack, aiming to penetrate through a weak point or to turn the enemy's flank. Once in the rear, French forces would envelop and encircle the enemy, thereby dislocating him physically and demoralizing him psychologically. This accomplished, French forces would fan out in pursuit, isolating enemy units and destroying them in detail.

These tactics were especially effective in Napoleon's three stunning victories in 1805–1806 against the Austrians at Ulm, the Russians at Austerlitz, and the Prussians at Jena-Auerstadt. Austerlitz is typical. There the French Army, not yet recovered from Ulm only a few weeks before, fought outnumbered by about 1.25:1 (85,000 vs. 67,000) and at the end of a long supply line. Expecting victory, the Russian Army attacked, its forces linearly dispersed across a 7-mile wide front. Napoleon lured the Russians into committing additional forces on their left by feigning retreat, and launched an offensive feint on the Russian right that led the Russians to shift units there. Quickly concentrating, the French Army then launched a massed assault against the weakened enemy center, throwing the Russian Army onto the defensive in a vulnerable linear array. The French promptly broke through,

[1]See Chandler, 1978.

encircled many Russian forces and destroyed others in a bloody pursuit. When the dust had settled, the Russians had suffered 27,000 casualties, the French 8,000. More important, the Russian Army was so badly mauled and demoralized that it no longer was capable of fighting.[2]

As a result of victories like Austerlitz, the French Army reigned supreme for several years, but the pendulum slowly began swinging backward as *La Grande Armée* began losing its élan and France's opponents developed countertactics. For a time, the French continued winning; Eylau, Friedland, Wagram, and Borodino are examples. But with decisive breakthroughs and envelopments less easily gained, these French victories often were more closely contested, more costly, and less decisive. Then came defeat. Initially the defeats that Napoleon began to suffer after the retreat from Moscow in 1812 were not a product of ineffective strategy. His opponents banded together to invade France with a vastly larger army that wore *La Grande Armée* down. His final defeat at Waterloo in 1815 was a parity battle (70,000 troops apiece) in which Napoleon's vaunted offensive schemes failed him.

Expecting victory and disdainful of the British, Napoleon launched a succession of concentrated attacks against Wellington's linear array, but his assaults were poorly coordinated and improperly executed. Equally important, the British were densely deployed on a three-mile front, benefited from favorable terrain, and showed good tactical skill. Every French assault consequently was rebuffed. Near the end of the day, the British were reinforced by Prussian troops. Wellington then counterattacked against the demoralized French and quicky routed them. The battle ended with both sides losing about 25,000 troops. But the French Army, shattered, retreated in disarray. Shortly thereafter Napoleon was toppled from power and went into final exile.

Waterloo presaged a similar failure of breakthrough tactics in the next major Western war, the American Civil War of 1861–1864. This conflict saw a great deal of cross-country travel as the two armies regularly marched long distances to engage each other. The battles themselves typically took the classic Napoleonic form in which the attacking side concentrated in an effort to punch through the defender's linear array. This particularly was the case along the eastern seaboard, where the Union Army of the Potomac fought against the Confederate Army of Virginia. During the war's first two years, when Confederate forces were still strong, many of these battles were fought at parity, or something like it.

The outcome was a steady stream of victories for the defender, always less dramatic than Waterloo but decisive enough to compel the battered attacker to break off and retreat. Early on, most of these victories fell to the Confederate side, which repulsed the Union Army at Fredericksburg and Chancellorsville. The Union side turned the tables at Antietam and Gettysburg. In striking terms, Gettysburg resembled Waterloo in form and content. The defending armies were similarly sized, with 75,000 Confederate troops under Lee attacking 90,000 Union soldiers. Over three days, the Confederates launched repeated but clumsily executed assaults against the Union Army's left, right, and center. The Union forces effectively met these assaults with a densely deployed linear array that enjoyed favorable terrain and a convex line that permitted rapid counter-concentration of reserves. Both sides suffered about 25,000 casualties; but the Confederate Army, rebuffed and badly mauled, retreated from the battlefield, never again mounting the strength to attack.[3]

The North ultimately won the Civil War, but not because of any superiority in strategy and tactics. It simply marshalled its larger industrial, material, and manpower resources to

[2]Chandler, 1978, pp. 413–442.

[3]McPherson, 1988.

wear down the South in a prolonged attrition campaign during 1864–1865. The interesting facet of this war is not that the stronger side ultimately won, but rather that when parity battles were fought early in the conflict, the defending army so often was victorious. It was this phenomenon that caused the war to drag on so long. What were the causes? Early in the war, prevailing sentiment held that the Confederate Army had better soldiers and leaders, but Gettysburg showed that other factors, affecting both sides, were at work.

One cause was that officers on both sides had learned the lesson of the Napoleonic Wars; defensive tactics consequently were more sophisticated and adaptive. Defending armies took care not to allow themselves to be lured into extending their fronts too far or otherwise sacrificing the advantage of mass. Defensive positions were selected with a careful eye to the terrain, while reserves normally were withheld from the frontal wall and kept in position to counter-concentrate quickly. Also, technology helped the defense. Artillery and infantry weapons were more lethal and thereby better able to destroy enemy columns as they approached. In contrast to the Napoleonic era, infantry rifles now had sufficient range to prevent the attacker from bringing his artillery close enough to punch holes in the defender's line. With the defender's line still intact and able to return coordinated fire, attacking formations normally suffered heavily as they approached the defender and were left too weak to win the close battle. Simply stated, the offense found itself ground down by the defender's combination of mass and firepower.[4]

By the time of the Franco-Prussian War of 1870, the pendulum had begun swinging back in favor of the offense. In this war, a German force of 384,000 soldiers attacked a French force that was in the process of deploying defensively into Alsace-Lorraine. Mobilizing more slowly, the French Army initially began the war with only 200,000 deployed troops, eventually building to 300,000. The German Army decisively defeated it after a turbulent two-month battle in which German organizational efficiency held sway over a series of tactical foul-ups on both sides.

The fight was waged over a battlefield 150 kilometers wide, marking the emergence of the large battlefield that has characterized Europe's 20th century wars. The growing range and lethality of modern weaponry had led German and French commanders to disperse their forces far beyond the narrow, densely packed formations of the Civil War. As a result, forces on both sides were spread out enough to permit tactical maneuver against one another. The Germans advanced with forces concentrated on their left wing, primed for a flanking and wheeling operation around the outnumbered French right. A complicated maneuver battle of stroke and counterstroke ensued. In the end, the Germans successfully broke apart the French defenses. The French Army's left wing was isolated in the town of Metz; the French right wing was enveloped, cut off, driven backward to the Belgian boder, and compelled to retreat. The conflict dragged on another six months, during which the Germans laid siege to Paris and Metz, and the French eventually surrendered.[5]

The French Army's decisive defeat was largely a product of the German Army's emergence under von Moltke and Bismarck as a first-class fighting force. Capable of mobilizing rapidly and deploying quickly on Germany's modern rail system, the German Army had a well-developed command and control system, combat organization, and logistic support structure. It also ranked high in technical planning, doctrine, training, and leadership. These assets helped it to perform coordinated, large-scale field operations on a battlefield in which forces operated well beyond the eyesight of their top commanders. The German Army's

[4]McPherson, 1988.

[5]Jones, 1987.

performance in these maneuvers was far from perfect; often junior officers had to alter established plans. Nonetheless, the German Army was able both to seize the initiative and to maneuver more quickly and effectively than the French, who let several counterattack opportunities slip away.

Aiding the German Army were modern rifles and long-range artillery tubes that also gave it a major advantage in firepower against an enemy that was not similarly armed and enjoyed no offsetting edge elsewhere. Even so, the war's initial battles were closely contested, and the Germans suffered some reversals. Their many advantages, however, were sufficient to enable the Germans eventually to overpower a French Army that, while fighting enthusiastically, did not meet its opponent's standards. The Franco-Prussian War of 1870 thus showed that operational maneuver had returned to the battlefield in ways enabling a skilled army, supported by modern weapons, to take the offensive in the expectation of a quick victory.

Germany's victory in this war gave rise to military reform efforts across Europe aimed at reconconfiguring national armies there to meet modern organizational and technical standards. It also contributed to the widespread belief, from 1870 onward, that the offense once again reigned supreme on the battlefield. The artillery and infantry weapons that later were to dominate World War I did not alter this belief. These weapons were seen as aiding the attacker more than the defender by providing the firepower needed to overpower prepared positions. This holds true even for the machine gun, which made its appearance in the late 1800s and, when initially used in small conflicts, seemed to help the attacker. As the new century dawned, armies across Europe increasingly oriented their forces, plans, and weapons toward the task of waging decisive offensive warfare. Defensive concepts were not ignored, but they primarily checkmated the enemy's maneuvers, thereby setting him up for a fatal counterthrust.[6]

This complicated 19th century history, which saw the pendulum swing first to the offense, then to the defense, and then back to the offense, set the stage for World Wars I and II. The previous century's wars, driven by the ebb and flow of doctrine and organization as well as weapons, had often confounded military experts by producing surprising, counterintuitive results. The 20th century's two world wars in Europe were to be no exception. World War I, which many originally had expected to be an offense-dominated affair, evolved into a prolonged attrition slugfest behind static defense positions. Examples of offensive failures to prosecute breakthrough campaigns against a linear defense are drawn from this war. By contrast, World War II, which experts wrongly predicted would deteriorate into another exercise in trench warfare, turned into an explosive war of offensive operations emphasizing theaterwide maneuver. The examples of offensive success and mixed results are drawn from this conflict.

DEFENSIVE SUCCESSES IN WORLD WAR I

In its early stages, World War I seemed unlikely to provide a massive case study in how a linear defense can rule the battlefield. Reflecting military doctrine at the time, contemporary observers believed that any war between Germany and France would be settled quickly, by maneuver. When invading German forces came pouring into Belgium and France in August 1914, setting off a titanic struggle between nearly 100 divisions and 1.5 million soldiers on both sides, initial events seemed likely to validate this prediction.

[6]Jones, 1987. See also Griffith, 1981.

Beyond doubt, Germany's military strategy was offensive and maneuver-oriented. Reflecting the desire to avoid a two-front war, Germany's "Schlieffen Plan" called for an offensive campaign aimed at knocking the French out within six weeks, before Russia could mobilize its own forces. On the left, two German armies were to engage and pin down the bulk of the French Army, which was concentrated around Metz in Lorraine. On the right, meanwhile, five German armies rapidly advanced through Belgium with the goal of turning southwest, attacking the rear of the French Army east of Paris, and destroying it. The French deployed their seven armies (supported by the British Expeditionary Force) in a forward-moving linear array along a 400 kilometer front. Their forces primarily concentrated in the center and were guided by plans to advance into Germany through Alsace-Lorraine, thereby striking at the hinge of the German offensive. As the interaction between these plans became clear, the outcome seemed likely to turn on which side could carry out its offensive design at the expense of the other.

The German Army, whose initial march swept 300 kilometers across Belgium and northern France in only 14 days, nearly accomplished the task. But 40 miles north of Paris, the German drive slowed because of logistic constraints, destroyed bridges, and growing allied resistance. Meanwhile the French Army, its drive through Alsace-Lorraine blocked, began quickly redeploying forces westward by taking advantage of railroads in the area. A subsequent British-French counterattack on the Marne threw the Germans 70 kilometers back to the Aisne River. Their frontal advance blocked, the Germans began moving in a northeasterly direction in an effort to turn the allied left flank. With the allies also trying to outflank the Germans, the two sides became engaged in a race that soon ended at the English Channel. With any prospect of a flanking attack now eliminated, the two sides began digging in with the goal of preparing strong defenses along their extended lines that would allow them safely to concentrate somewhere and attack. With this development, World War I abruptly changed course and became a contest of trench warfare.

The five battles examined here, which were among the largest and most notorious engagements of the war's frustrating period of static trench warfare, grew out of efforts by both sides over the next four years to breach the defenses of the other. They all were fought in the hotly contested area between Ypres and Compiègne, a 150 kilometer sector that saw some of the fiercest fighting and heaviest casualties of World War I (see Fig. 32). Because both armies engaged in either attacks or costly counterattacks, losses were evenly balanced and very high on both sides.[7]

- *First battle of Ypres* (October–November 1914). A force of 14 German divisions attacked along a 20 kilometer front. Opposing them were six frontally committed British and French divisions, quickly reinforced by three more divisions drawn from neighboring sectors. Over a two-week period, the Germans advanced only about 5 kilometers at a loss of 135,000 casualties. The allies also lost heavily.
- *Second battle of Ypres* (April–May 1915). Again at Ypres five months later, 15 German divisions attacked along a 25 kilometer front against eight allied divisions. Employing poison gas, the Germans opened a four mile gap in the allied line, but were unable to bring up reserves fast enough to exploit the opportunity, and the allies promptly assembled another eight divisions to heal the breach. Over a three-week period, the Germans advanced only about five kilometers before bogging down; again massive losses were inflicted on both sides.

[7]Jones, 1987. See also Natkiel, 1982.

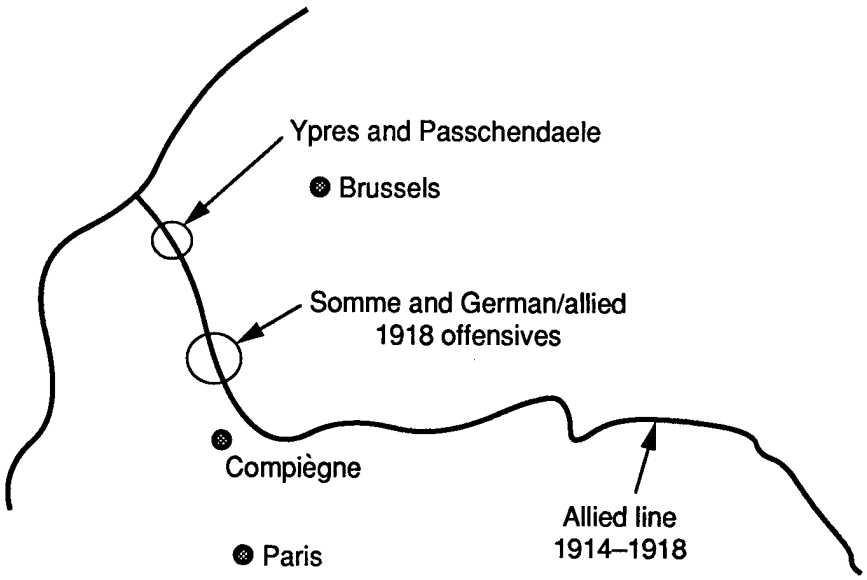

Fig. 32—Trench Warfare in World War I, 1914–1918

- *The Somme* (July–November 1916). In this bloody battle, 22 British and French divisions advanced along a 45 kilometer front. Opposing them were ten German divisions, quickly joined by seven reserve units. A week-long preliminary allied artillery barrage was intended to suppress the German defenders but failed. As a result, advancing British and French forces were slaughtered by German rifle, machine gun, and artillery fire. Several followup allied assaults yielded a total advance of 10 kilometers at a loss of 600,000 allied casualties. German losses also were heavy.
- *Passchendaele* (July–November 1917). Some 15 British, Australian, and French divisions attacked along a 25 kilometer front against 11 well-entrenched German divisions. Both sides committed additional forces as the battle dragged on. Over a 13-week period, the allies advanced 10 kilometers; casualties were about 250,000 for both sides.
- *Germany's 1918 offensive* (March–July 1918). By this stage of the war, Russia, now under Bolshevik rule, had dropped out. Germany, freed from what previously had been a two-front war, now was able to begin transferring forces to the western front. With this added strength, the Germans mounted a major attack aimed at defeating the allies before American forces, already beginning to arrive, could deploy in strength. The result was this massive and climactic battle, German's last offensive of the war. Some 80 German divisions advanced along a 200 kilometer front. Allied forces initially were outnumbered by 2:1, but as reserves became available, the ratio came close to 1:1. Three sequential German drives yielded a total advance of 40 kilometers. For a time, the Germans appeared on the verge of overpowering the allies, but the British and French brought sufficient forces into the area to stop the advance.

With the German Army exhausted and demoralized, the allied forces, gaining increased strength from U.S. reinforcements, concentrated that August in the area between Amiens and Compiègne. Their attack yielded a 20 kilometer advance. Then in September, 160 allied divisions launched a massive drive in that area against 100 German defending units. For the first time in four years, the allies gained appreciable ground: Over a period of six weeks, their drive carried 100 kilometers into Belgium. In the face of this setback, Germany, its army weakened and its domestic support collapsing, sued for peace. The war ended shortly thereafter in Germany's defeat.

All five of these battles involved a concerted effort by a strong attacker to breach a well-prepared positional defense. All five failed by a wide margin and at great cost. In each case, the attacker not only fell short of a major breakthrough but also failed to make any substantial gain in territory. What were the causes of these defeats, and what do they suggest about the viability of linear defense in general?

One commonly accepted theory cites three mutually reinforcing factors behind the steady stream of defensive successes in World War I: the elaborate German and allied trench networks, the nature of military technology at the time, and the unimaginative tactics by senior commanders on both sides. This explanation basically holds that the defense was so well-entrenched and heavily armed that it was invulnerable to the offense's blows, regardless of how heavily concentrated or well-prepared. Offensive strategy primarily relied on artillery fire to blast holes in the defender's line, through which waves of attacking infantry were then to advance. But World War I artillery lacked the explosive power, accuracy, specialized munitions, and logistic support needed to suppress a dug-in defense, with trench lines often laid out in depth. About all a lengthy artillery barrage accomplished was to alert the defender that an attack was imminent, thereby sacrificing the advantage of surprise and tearing up the ground so badly that attacking forces could not advance rapidly.

Regardless of how intense the preparatory artillery barrage, this explanation holds, defending troops were able to emerge pretty much intact. Attacking infantry, advancing slowly over muddy and rough terrain, consequently encountered withering fire from the defender's magazine rifles, machine guns, and artillery. The outcome normally was a slaughter that gained little, if any, ground. And when small penetrations did occur, the defender was able to mount a counterattack to repel the invaders. Only armor presumably could have changed this situation. However, it was not available until late in the war, and its value was not fully appreciated even then. Offense thus ceased to be a viable option. Yet the two sides kept resorting to frontal assaults. The motive presumably was political considerations, reinforced by an unflinching faith among military commanders in the predominance of the offense.

This explanation clearly has considerable validity, but there are reasons for not letting it stand alone or unquestioned. With respect to motives on both sides, perhaps the German government did not have any coherent goal beyond victory for its own sake, but it was not blind to military realities. Recognizing in 1915 that the western front had stalemated, it largely swung over to the defense to pursue the offense on the eastern front, where the situation was more fluid. The Germans did not return to the offense in the west until 1918, after the Russians had been defeated. During the intervening years, it was the allied governments who principally pursued the offense, and they had a valid reason for doing so. Since German forces were occupying Belgian and French territory, the allies could hardly have afforded to accept their situation passively. Constant offensive pressure was necessary either to eject the Germans or to compel them to come to terms.

With respect to military doctrine, senior commanders clearly did believe in the offensive, but they were hardly impervious to the appeals of more subtle tactics than frontal assaults against prepared fortifications. These, after all, were the same officers who had boldly conceived the sweeping maneuver operations that opened the war. The fact that they repeatedly resorted to frontal attacks was less a product of blind ignorance to the alternatives than a recognition that little else, short of inactivity, was in the cards. They clearly did overestimate the ability of massed attacks to penetrate prepared defenses. Their goal in these attacks was not to perpetuate attrition war for its own sake; at a minimum, they wanted to wear down the enemy and break his morale. Ideally, they wanted to open the battlefield for maneuver by breaching the enemy's frontal wall, gaining a breakthrough, and advancing into the open rear areas. In the end, they failed. But all they needed was one success, and their judgment at the time was that the game was worth the candle.

Even after the war had settled down into fixed positions, their decisions on exactly where to attack had an operational logic to them. The bulk of German and allied attacks were conducted between Ypres and Compiègne, where the allied line bulged inward in concave fashion. This was the one sector where a German breakthrough might have unhinged the entire allied defense. The allies, in turn, had powerful reasons to launch attacks of their own in the hope that even limited advances might straighten out this salient. As a result, the two sides engaged in a bitter, but not unreasoned, tug-of-war over this sector.

Over the years, tactics and weapons also evolved a great deal, a sign of intellectual ferment aimed at breaking the deadlock. The crude artillery barrages and massed infantry assaults of the early years went nowhere. As the war dragged on, the development of better artillery techniques, air reconnaissance, and improved small-arms weapons began improving the attacker's lot. Also, both sides became more skilled at using infiltration tactics to bypass defensive positions that, by this time, were becoming elastic networks of interlaced strongpoints rather than linear trenches. As a result, by 1917 attacks began penetrating deeper than before and coming closer to success. The primary reason why they continued to fail was not that the defender's positions were impregnable, but that the attacker's lead echelons lacked the radios to call exploitation forces forward to the point of an impending breakthrough, and these forces lacked the mobility to get there fast enough once they were informed. The defender, enjoying greater mobility from rail lines in the rear, consequently always was given sufficient time to read the situation and bring reserves to bear. Even so, both the Germans and the allies gained considerably greater ground in 1918 than had been the case earlier.[8]

These considerations suggest that, early attacks aside, the later battles of World War I were more closely contested than is realized today. Was it, then, the defender's inherent advantage in firepower, and its disproportionately lethal effects, that swung the balance? Perhaps so. However, the fact that defending forces in these battles normally also suffered heavy casualties is one indication that the attrition process was not one-sided. To some degree, the military theorists who believed that World War I's weapons could work to the attacker's advantage were not entirely misguided.

Any attempt to explain World War I's dynamics must also take into account another factor, one so fundamental that it is easily overlooked: the enormous size of the armies during this war in relation to the terrain on which they were fighting. Both sides began the war with nearly 100 divisions, and they quickly began enlarging their forces by fully mobilizing

[8]Jones 1987; and Natkiel, 1982.

their national populations and industries. By the end of World War I, each side deployed 175–200 well-armed combat divisions along a front line that was only 550 kilometers long. With so many units, each side had the luxury of achieving very high troop densities at critical places, while neither dangerously thinning its line elsewhere nor denuding itself of reserves.

In the battles examined here, division frontages in threatened sectors tended to be as low as 2–3 kilometers, not much wider than the frontages regularly achieved by attacking forces. This enabled each division not only to form an initial front line but also to deploy in depth. Defending units consequently were well situated for standing firm in the face of massed infantry assaults, slowing them down, and inflicting high losses on them. Meanwhile, the defender, using his rail lines in the rear, was able to draw on other forces deployed nearby and to rush them to the point of contact. With these forces available as instruments of counter-concentration, the defender was seldom vulnerable even when gaps temporarily opened in his line.

Because of the defender's consistently high troop density all along the front, advance concentration and surprise mattered for less than has been the case in many other wars. With the attacker unable to gain a big local advantage by swiftly and secretly massing, World War I's breakthrough battles were fought under conditions in which a 2:1 ratio, maintained only temporarily, seems to have been a normal upper limit on the attacker's initial advantage. Shortly after that, the ratio normally converged toward 1:1 as the defender began responding by redeploying his forces from elsewhere. This is a far cry from the sustained 3–5:1 local advantage that commonly is cited as an appropriate goal for expecting local success. To compound matters further, the defending forces normally dominated the terrain to such an extent that open, inviting avenues of advance did not readily beckon. They had to be created by either forceful frontal assault or penetration tactics. Even at the war's end, this task proved to be either too costly or too difficult for slow-moving infantry to accomplish.

World War I thus was a "stable" conflict in the sense that neither side was able to use clever tactics and strategems to turn the operational tables on the other. But this outcome was a product of more than firepower and prepared positions alone. Mass also mattered, and in a dominating way. Had the forces on both sides been much smaller and more thinly deployed, defensive firepower and field fortifications might well have played a less influential, less paralyzing role. In this event, the war might well have become the more explosive, maneuver-dominated affair that its architects originally envisioned.

ATTACKER SUCCESSES IN WORLD WAR II

Whereas World War I saw the defense reign supreme, the offense returned to prominence in World War II. The battlefield opened and theater-level maneuvers came back into fashion. Breakthrough operations followed by flanking attacks and rear area exploitations reclaimed an important role in military strategy, and linear defense worked less effectively. This particularly was the case during the war's first two years, when the German Army rode its blitzkrieg doctrine to a steady stream of victories against equally large but outclassed opponents. Two German victories will be examined here, the Battle of France in 1940 and, in less detail, the Wehrmacht's invasion of Russia (Barbarossa) in 1941. These battles provide classic examples of how the theory of maneuver war has been successfully put into practice against an exposed linear array that left the defender's forces caught out of position.

When Hitler invaded France and the Low Countries in May 1940, the Wehrmacht on paper was no stronger than the combined strength of the allied forces opposing it. In total, the Germans fielded 136 divisions; 94 were committed initially and the remaining 42 were withheld as reserves. By comparison, the allies fielded 142 divisions, 104 of them French. Some 116 allied divisions were committed to the forward battle and the other 22 were held in reserve. In maneuver units and manpower, then, this was a parity fight.

Nor were there appreciable differences in weapons. Individually, allied tanks, artillery, aircraft, and other weapons were as good as their German counterparts, sometimes better. The Germans had a slight numerical edge in tanks (3000 vs. 2700) and a wider lead in combat aircraft (3000 vs. 2000); this imbalance in aircraft narrowed somewhat as British fighters started deploying to the continent when the invasion got underway. German armor and airpower were better organized for conducting a maneuver war, but organization aside, the German numerical edge in weapons fell well short of tipping the military balance decisively. Measured in purely static terms (WEI/WUV scores), the two armies were equal.

Yet the Wehrmacht swept to a complete victory in the short span of six weeks, thereby fulfilling, on the same time schedule, the wishes of its predecessor in August 1914. Attacking on May 10, the Germans, employing daring airborne assaults to pave the way, overpowered the ten-division Dutch Army in four days. Then on May 14, German forces advancing through the Ardennes forest to the south, broke through French defenses between Namur and Sedan. After crossing the Meuse River, German armor spearheads broke out into the open rear areas, turned westward, and sped 250 kilometers to the English Channel in five short days. This maneuver cut off 60 French, British, and Belgian divisions that earlier had deployed along the Dyle River some 100–150 kilometers to the north. The subsequent British evacuation at Dunkirk saved 330,000 allied troops, but nearly all their equipment was left behind. By May 31, consequently, this entire allied force had been eliminated as a factor on the battlefield.[9]

The Wehrmacht then turned southward to attack the remaining, now outnumbered, 72 divisions of the French Army, most of which originally had been deployed on the Maginot Line in Alsace-Lorraine. On June 4 the Germans broke through hastily prepared French defenses on the Weygand Line, exploded into the French countryside, fanned out, and overran French forces in their way. Shortly thereafter, three French armies were trapped behind the Maginot Line. On June 22, the French government surrendered, its army routed and its spirit broken. An allied army of nearly two million men had been decisively beaten, at a cost of only 140,000 casualties to the Germans.[10]

How was this victory achieved, and why did the allied defenses collapse so totally? The Maginot Line often is blamed for France's defeat, but this line of fortifications, built to guard France's eastern border with Germany from Luxembourg to the Swiss border, played little role in the battle. It was originally designed to block enemy advances on the open terrain that the Germans had used in 1870 to defeat France. The hope was that it would free large French forces to guard the northern border, which the Germans had used in 1914. For the most part, the Maginot Line did its job. Although it was penetrated in some places, the Germans gained no fatal breakthroughs there. The important battles were fought to the north, especially between Sedan and the English Channel, where few fortifications had been built and the allied armies found themselves engaged in an open war of maneuver.

---

[9]See Liddell Hart, 1967, for a political-military analysis of German strategy in the 1940 campaign. See also Fuller, 1966. For an appraisal of European military doctrine in the 1930s, see Posen, 1984.

[10]See Natkiel, 1982.

Another argument is that the French and British brought a "defensive mentality" to the battle, born of their World War I experience, leaving them vulnerable in World War II. Although this was true to some degree, it is not a fully accurate explanation of the allied mindset as the battle opened. The allies had seven months between Poland's defeat in September 1939 and the German invasion of France the following spring. The fact that they did not take advantage of the opportunity to recreate a network of trenches along France's northern border is an indicator that they recognized the times were changing. Also, the allies did, in fact, maneuver during the battle. As the Wehrmacht swept into the Netherlands, French and British forces, pivoting on the Sedan axis, quickly marched northward to establish themselves on the Dyle River, hardly the behavior of an army locked into fixed positions.

On paper, the Dyle maneuver had attractive features, especially for an army that anticipated a German assault from the north. The allies' defense line north of the Maginot Line, which would have been 300 kilometers long had their forces remained in France, was now shortened to about 240 kilometers. Also, French and British forces were now joined with the Belgian Army's 18 divisions, thereby increasing the allied posture by nearly 50 percent. With 56 divisions defending a line of 240 kilometers, the allies now had enough density, by doctrinal standards of the day, to contemplate a stiff linear defense. Divisonal frontages averaged only about 4 kilometers apiece, well below the allowable limit of 8–10 kilometers. The allies had no major network of trenches and prepared positions to help protect them. But such terrain features as Antwerp, the Dyle River, the Meuse River, and the Ardennes all now screened a German advance, and the allies had gained greater strategic depth in which to operate. Ironically, it was this very penchant for early and opportunistic maneuver, rather than any fixation on positional warfare and fortifications, that helped get them into trouble.[11]

Maneuver aside, one genuinely important factor behind the allies' defeat was that the German Army, on a unit-by-unit basis, simply fought better than its opponents. German divisions were better trained, better led, and better able to execute their doctrine. During the interwar years, German military leaders, refusing to accept World War I's verdict that modern defenses were impregnable, had resurrected the swift, decisive theater offensive. Sensing the potential importance of armor and airpower, the Germans organized their tanks and combat aircraft so that they could play a leading role in the new Wehrmacht's offensive doctrine. Among a host of innovations, they brought their tanks together to form ten tank divisions that were used to spearhead the Ardennes breakthrough. They also harnessed their combat aircraft to support their armored spearheads by providing air cover along with attacks against enemy positions. Above all, the Germans mastered the art of combined arms warfare. Armor, motorized infantry, foot infantry, artillery, airpower, and motorized logistics were forged together into a cohesive posture capable of prosecuting a swift-moving offensive.

By contrast, the allies had not fully anticipated the changing nature of warfare. They had procured large numbers of tanks and airplanes, but they failed to organize them properly. They continued to anchor their posture on foot infantry and artillery. Armor and motorized infantry were scattered across the posture in small doses rather than concentrated to form large, mobile forces. Airpower was not closely tied to the ground battle. All this yielded a large and seemingly powerful posture that was not highly capable of combined arms operations. To complicate matters further, allied troops were less well-trained and their leaders were less professionally skillful than the Germans. Even had the battle followed a more predictable course, these problems would have left the allies vulnerable.

[11]Jones, 1987.

It was the interaction of the two sides' strategies and the maneuver dynamics thereby unleashed that produced the allied defeat. Once the Germans had selected the Ardennes and the allied center as the main point of attack, they concentrated their forces there. Figure 33 shows that the Wehrmacht's battle plan allocated 29 divisions on its northern flank under Army Group B, which swung through Holland and northern Belgium to reach the Dyle River north of Namur, opposite Holland. In the center, opposite the Ardennes, were 45 divisions under Army Group A. To the south, opposite the Maginot Line, were 19 divisions under Army Group C. Some 42 divisions were held as a strategic reserve. The German operational plan hoped to dislocate the allied defense by lifting it off its hinges. The idea was to lure the

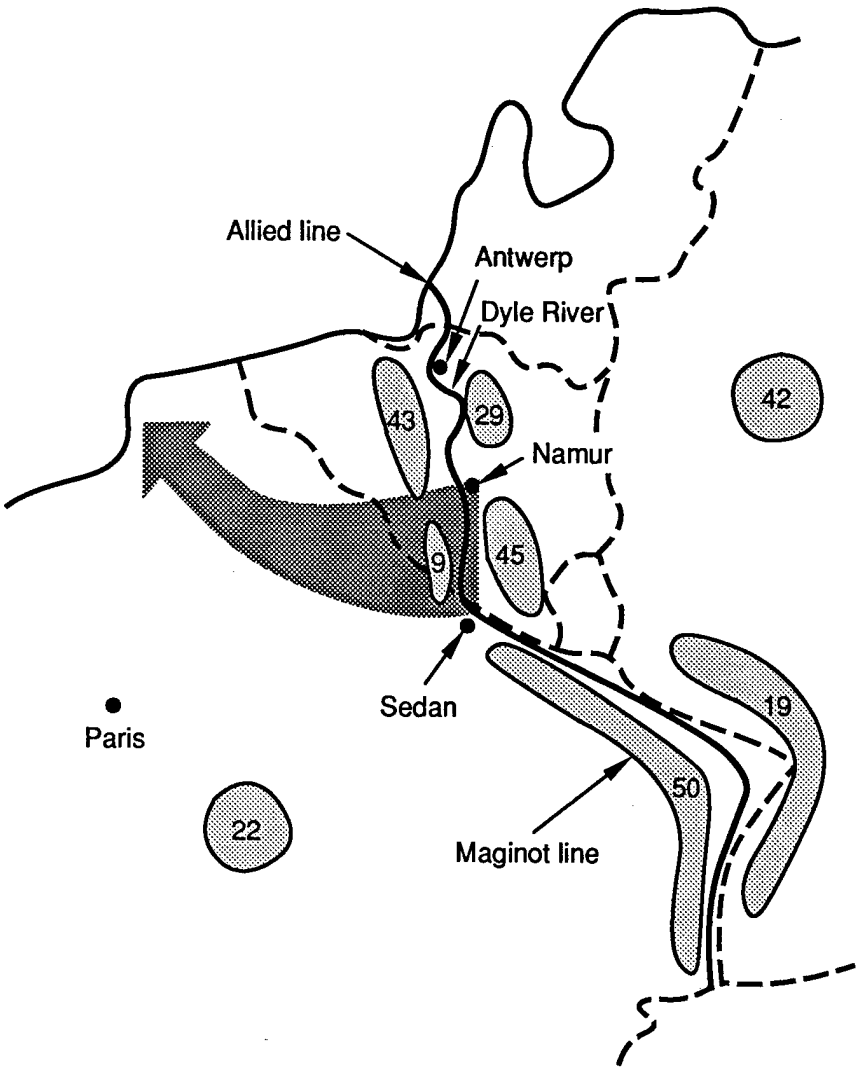

Fig. 33—German and Allied Force Deployments, May 12, 1940

allied left and right wings forward, thereby exposing them to rearward envelopment once a German breakthrough at the center had been gained.

The allies played directly into the Germans' hands by deploying in quite different fashion, choosing strength on each flank at the cost of a weak center. Also, 85 percent of allied forces were deployed along the entire front line, with only 22 divisions held as a strategic reserve. As a result, the allies held nearly a 1.5:1 advantage on the north from Namur to Antwerp and a 2.6:1 edge on the south, opposite the Maginot Line. But in the critical center from Namur to Sedan, the allies found themselves outnumbered by 3:1 in maneuver units and by about 5:1 in tanks, with few reserves immediately available. This locally adverse ratio was to prove decisive.

The area between Namur and Sedan, guarded principally by the French 9th Army, pointed like a dagger at the heart of the allied defense. A German breakthrough there could open the gates for a westward drive to cut off the allied forces on the Dyle and sever their supply lines. Alternatively, a southward turning movement aimed at entrapping French forces on the Maginot Line also was possible. Allied commanders were not unaware of this vulnerability but attached higher priority to bolstering the Dyle Line and the Maginot Line to the south. Consequently they deployed only seven infantry divisions and some lightly armed cavalry reconnaissance units to this 100 kilometer sector. Depending upon how much weight is given to the cavalry units, the resulting force-to-space ratio yielded an average frontage of 10–12 kilometers per infantry division. This was at the outer limits of doctrinal tolerance and over twice the distance covered by most divisions on the Dyle. Equally important, all of these divisions had to be committed to the front line to achieve even this density. The 9th Army thus had little depth; the nearest reserves were several miles away, under command of the French 1st Army to the north and the 2d Army to the south.

The French commander, General Gamelin, based the 9th Army's tenuous deployment pattern on two fateful judgments. First, he believed that the Germans were unlikely to send large armored forces through the rugged Ardennes. Second, he judged that even if the Germans did concentrate there, the forested terrain and local French forces would delay them for several days, allowing time for allied reserves to converge on the scene and contain the attack. Gamelin was proved wrong on both counts. The reasons why his second judgment was faulty help explain how the Germans were able to gain and then exploit their breakthrough.

Driving through the Ardennes in secrecy and much faster than the French expected, the Germans reached the Meuse River on May 12. But instead of pausing 4–5 days to bring up their artillery, infantry, and logistic support as expected, the Germans immediately crossed the Meuse at Sedan and Dinant. Within a day, a sufficient bridgehead had been developed to begin bringing large numbers of tanks across. Shortly thereafter, armored spearheads, supported by tactical air strikes substituting for artillery, began driving westward in an effort to breach the French defenses.

On paper, the French forces there should have been able at least to delay the Germans. The French front line, while less dense than desirable, was hardly paper thin. Also, the battle was fought against only the nine-division lead echelon of the German column. The manpower ratio thus was close to 1:1. The Germans had a big edge in tanks and aircraft, but the French had more artillery pieces since the Germans attacked before their own artillery could arrive in strength. The firepower ratio therefore was not itself unmanageably high and the French had the advantage of being at least partially dug in in some places. The Germans, however, concentrated their nine divisions at four points along the front, thus gaining 2–3:1 advantages that brought French units there under considerable pressure.

The outcome hinged on whether these French units could fight effectively against the combat-experienced Germans. A few units did fight well; one well-trained French division, deployed in prepared positions north of Sedan, initially rebuffed the German river crossing. But overall the French performed poorly. Part of the reason is that some French battalions had not yet fully arrived at the scene, and most of the battalions already there were not yet well entrenched, nor were they properly armed. Although well endowed with artillery and small arms, they lacked the number of tanks and antitank weapons needed to contest enemy armored divisions. To compound their problems, none had combat experience and several were poorly trained and led. These drawbacks left them ill-prepared for the air bombardment and massed tank assaults that they faced from May 13 onward.

At Sedan, German General Guderian advanced with over 800 tanks supported by Stuka dive bombers. The air attacks effectively suppressed the French artillery, thereby leaving the ill-equipped infantry exposed to Guderian's tanks. In the face of the German onslaught, the two French divisions there disintegrated, many of their troops and officers fleeing in panic. With a path opened, Guderian immediately began racing westward. To the north at Dinant, General Rommel attacked with 500 tanks. The two French divisions there were better trained and fought hard at first, but they were guarding an unusually long 30 kilometer front that began expanding as the German bridgehead swelled. Soon German tank spearheads penetrated the thinly spread French battalions and infiltrated into the rear, encircling and destroying enemy positions along the way. What had started out as a classic set-piece battle was suddenly turning into a rout, with the French rapidly losing control.[12]

With their northern and southern flanks crumbling, the French tried to pull back all along the 9th Army's front and form a defense line to the rear. But with German armor advancing so fast, they were unable to move rapidly enough. Their line irreparably broken, the French counterattacked on May 15 with some reserves from the 1st and 2d Armies that had arrived at the scene the night before. An armored division struck against Rommel in the north, and another armored division and a motorized division attacked Guderian in the south. Pitched battles were fought in both places and the French tanks performed well. But the French soldiers operating them showed poor mastery of armor tactics, and the Germans quickly won both engagements. Pressing on, the Germans on May 16, already 60 kilometers beyond the Meuse, broke into the open countryside along the entire sector from Namur to Sedan.

The Germans' subsequent lightning drive to the English Channel could have been stopped had the allies been able either to maneuver forces into blocking position or, failing that, to strike hard against the German flank. The allies were able to do neither. A large number of French and British divisions theoretically were available for this purpose, but caught badly out of position and committed too slowly and inexpertly by allied commanders who were unprepared for the speed of the German advance, only a handful were able to reach the scene in time. This was too few to halt the German juggernaut of nine tank divisions and three motorized divisions with 35 infantry divisions trailing behind them.

The failure of allied forces to react more promptly gave the impression that a kind of paralysis had set into the allied command structure. It was not because senior French and British officers were blind to the risk that a German attack might come in an unexpected area, nor were they unaware that armor might be important or unschooled in the concept of rear-area maneuver. They formed an accurate picture of German ambitions fairly early and grasped the danger facing them, but their decisionmaking was sluggish, not by traditional

[12]Liddell Hart, 1969.

standards or in absolute terms, rather in relative terms. Relativity, however, is what counted.

Allied commanders were used to slower-moving infantry warfare and were unprepared for the fast tempo that German armor set both tactically and operationally. According to their doctrine the Meuse River crossing should have taken four days, the breakthrough 5–7 days, and an unopposed drive to the English Channel another 10–15 days. The Germans crossed the Meuse in one day, broke through in less than three days, and sped to the Channel in four days. All three phases of the German operation thus were completed in less than half the time the allies would have expected. Faced with this unprecedented tempo but not anticipating its implications, allied commanders regularly overestimated the time available to them. Consequently, they made their decisions to redeploy forces too slowly and then misallocated these forces by sending them to the wrong places. These failures in command interacted with the fact that allied reserves were badly maldeployed to respond quickly to an Ardennes offensive.

The result was a catastrophe that need not and would not have happened had the Germans themselves not been so daring and so efficient. Many German commanders originally had wanted a less risky northern attack guided by less ambitious goals and regarded the idea of attacking through the Ardennes with misgivings. Hitler had adopted the plan, originally proposed by General von Manstein, over their objections. Once the breakthrough occurred, the senior German hierarchy, nervous and fearing an allied flanking attack, twice slowed the advance to allow artillery, infantry, and logistic support to catch up. It was largely local Wehrmacht commanders, especially the audacious Guderian, who sensed the opportunity and insisted on continuing their rapid drive. Their boldness set the stage for the allied debacle that unfolded in the rear during May 16–20.

The French had some 20 divisions deployed on the northern boundary of the Maginot Line, about 80–100 kilometers from Sedan. Had these forces been committed when the Meuse was crossed and allied defenses began unravelling, they could have blocked the German breakthrough at its origins on May 16. But French commanders, fearing a German wheeling movement southward against the Maginot Line, hesitated. By the time the French realized that the German goal was the English Channel, these units were already too far behind to catch up. Some 20 theater reserve divisions in Central France then began moving forward, relying on rail to reach the French-Belgian border. The lead elements took four days for the journey and then had to detrain, assemble, and march forward. This might have been fast enough in World War I, but the Wehrmacht's mobility had narrowed the time window of modern warfare. The allied units arrived slowly and in a scattered fashion and were unable to launch any large attacks on the German column. Allied forces to the north, meanwhile, were so preoccupied with attacking German forces and weakened by the surrender of the Belgian Army that only a few managed to move south in time.

Unable to block the lead echelons of the German column, the allies did manage to launch two more armored counterattacks against the German flank along the way. On May 19, General Charles de Gaulle led his 4th armored division against the German flank near Laon, 75 kilometers from Sedan. Then on May 21, elements of three allied armored and motorized divisions attacked near Arras. These attacks slowed the German advance as sharp fighting occurred in both places. Once again, allied tanks performed well, but the Germans, far more proficient tactically, won. Shortly thereafter, they resumed their race westward.

In all, the allies brought seven armored divisions and five motorized divisions to the Battle of France, a posture that compared favorably with the Germans' ten tank and five motorized divisions. But the Germans employed their heavy divisions en masse, while the allies dispersed them. In their three counterattacks, the allies managed to get seven of their 12 heavy divisions into the battle. Had these divisions been committed en masse, they might have stopped the German spearhead of ten heavy divisions. Once German armor had reached the English Channel near Abbeville on May 20, it wheeled northward. By May 27 it had reached Calais and was followed shortly by large infantry forces. Although Dunkirk and the final German drive south were yet to come, the Battle of France, modern history's classic case of an attacker riding maneuver to victory, essentially was over.

Barbarossa

The Germans repeated their blitzkrieg doctrine a year later when Barbarossa was launched on June 22, 1941. Once again, the force ratio in maneuver units was 1:1 as both sides deployed about 135 combat divisions. This was virtually the same number that had fought the Battle of France. The Barbarossa battlefield, however, was much bigger. Whereas the front in France initially was only about 500 kilometers long, the Russian border from the Baltic to the Black Sea stretched 1500 kilometers. Moreover, the battlefield was much deeper. In France, the English Channel was located only 350 kilometers from the Ardennes. In the east, Moscow was located 900 kilometers from the Russian border. These vast distances changed the nature of the battle by opening up the terrain to offensive maneuver. They also made linear defense far more difficult by increasing the ratio of space to force. Barbarossa's outcome was driven primarily by the fact that the Germans made the most of this situation while the Russians fell victim to it.

The German operational plan followed the Wehrmacht's penchant for concentration and maneuver. The invasion force was divided into three Army Groups: North, Center, and South. Since Army Group North had only 26 divisions, the plan had the appearance of being oriented south. In reality, it was anchored in the north. Army Groups North and Center, with 77 divisions between them, were both deployed on the front's 500 kilometer northern sector from the Prippet Marshes to the Baltic Sea. Army Group South covered the remaining 1000 kilometers of front, with its forces concentrated somewhat south of the Prippet Marshes. This deployment permitted large German formations to drive toward Moscow and Leningrad, while secondary attacks were being launched into the Ukraine and toward Crimea. Figure 34 illustrates how Barbarossa unfolded between June 21 and the following December, when it ground to a halt at the gates of Moscow and Leningrad.[13]

Barbarossa unfolded as planned and initially met even Hitler's lofty expectations. At dawn on June 22, heavy German air attacks shattered Red Air Force units based nearby and German troops advanced into Russian territory up and down the long border. Facing them were 128 of the Red Army's divisions deployed forward in a linear array along the border. Although the Red Army tried to hold its ground, it was quickly pushed back and penetrated along nearly the entire front. Gaining multiple breakthroughs, German armored columns fanned out into the rear areas, encircling and trapping numerous Russian formations along the way. Meanwhile, the Luftwaffe supported the ground offensive not only by suppressing the Soviet air force, but also by bottling up Red Army units and disrupting their efforts to react. Attacking from East Prussia, most of Army Group North advanced toward Leningrad.

[13]See Seaton, 1971. See also Werth, 1964.

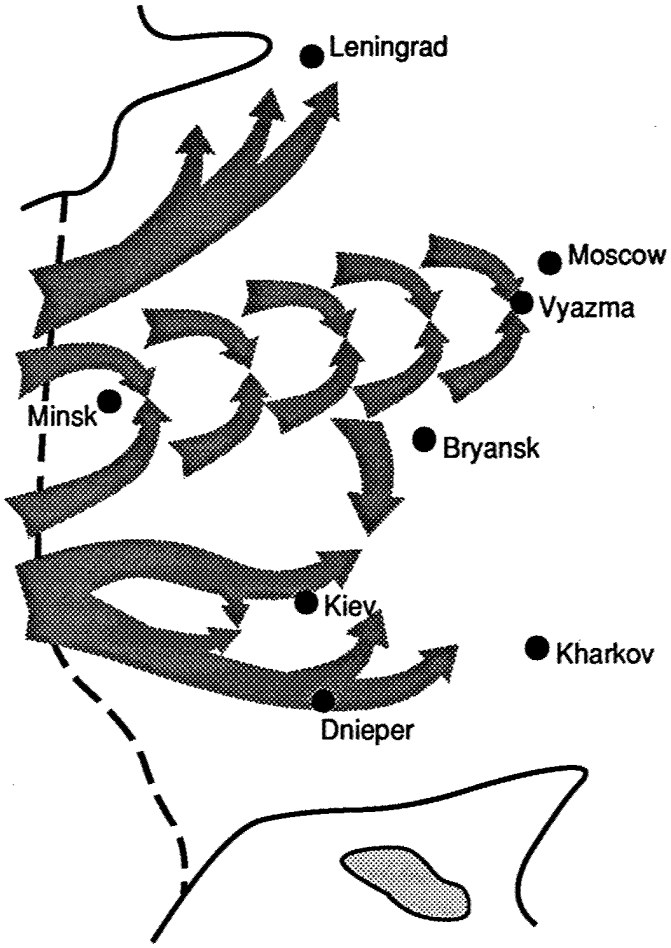

Fig. 34—Barbarossa, June–December 1941

Army Group Center, supported by elements of Army Group North, made a giant pincer movement that converged a week later at Minsk, 300 kilometers to the rear. Parts of two Soviet armies were caught in the pincer; over 300,000 Russian troops surrendered. Army Group South drove deep into the Ukraine toward Kiev, again capturing large numbers of Russian troops.

With the Red Army now in full but clumsy retreat, the Germans sped eastward. By July 9, the Wehrmacht had arrived at the vaunted Stalin Line, a lengthy but poorly prepared network of trenches, barriers, and prepared positions 500 kilometers deep into the USSR. The Germans quickly broke through, fighting a tough tank battle near Smolensk. In early October, two German pincers trapped one Russian army at Vyazma and another at Bryansk. By then, the Germans had advanced some 750 kilometers; Wehrmacht units already were taking Leningrad under siege and had drawn within 150 kilometers of Moscow. To the south meanwhile, ill-advised Soviet counterattacks led to further entrapments as the Germans seized Kiev, crossed the Dniepr River, and drew near Kharkov.

At this juncture, however, the German advance began slowing down. With its supply lines extended and the Russian terrain subjected to Stalin's scorched earth policy, the Wehrmacht was beginning to run short of ammunition, petrol, food, and other provisions. German logistic units were capable of transporting supplies by truck only to a distance of 300 kilometers. Railroads were needed in the rear areas to permit deeper drives. As it advanced eastward, the Wehrmacht twice had to slow down to allow German engineers to repair rail lines that retreating Soviet armies had destroyed. Even after these lines had been rebuilt, German railroad supply was not proficient. This problem took an increasing toll as the Wehrmacht approached Moscow.

Equally important, Soviet defenses had begun to stiffen. Although the Red Army had suffered huge losses, several units had escaped German armored pincers that, unlike the situation in France, had no ocean to pin the enemy against. To the aid of these retreating formations came a growing number of reinforcements. By early December, the Red Army, benefitting from the mobilization of manpower reserves and new weapons built by factories in the USSR's eastern districts, had grown from 11 to 30 separate armies. With over 300 divisions, it now had enough manuever units to establish a strong defense line, especially along the vital Leningrad-Moscow axis. Although Hitler committed 51 divisions to attack Moscow, their drive reached the capital's western suburbs but stalled there. With cold weather rapidly coming, Hitler called off the attack for the duration of the winter, thus bringing Barbarossa to a close.

Politically, Barbarossa fell short of its goal of seizing Leningrad and Moscow, a failure that presaged Hitler's ultimate defeat. But in purely military terms, Barbarossa was a major success. In only about four months, the Wehrmacht had driven 900 kilometers into Soviet territory along the entire breadth of the country, and it was still largely intact. Major losses had been inflicted on the Red Army, especially through encirclements that produced about one million surrenders. Counting these surrenders, Soviet casualties totalled three million, equal to the manpower that the Red Army had deployed at Barbarossa's onset. The Germans lost about 800,000, or about 25 percent of their original forces. Red Army losses doubtless would have run higher had the Red Army not engaged in a wholesale retreat once its front line had been destroyed. And had the Germans not diverted two armies from Army Group Center to support the drive to the south, Moscow itself might have been completely encircled and cut off.[14]

As in France, the Wehrmacht's success was owed heavily to its superior performance rather than any overwhelming preponderance in forces and armaments. Indeed, the Red Army deployed almost twice as many tank divisions as the Germans when Barbarossa began. The Wehrmacht simply was vastly better at combined arms operations. In addition to not grasping combined arms warfare, the Red Army showed errors in command and control similar to those of the French. Because Stalin had purged the Red Army's leadership in the 1930s and its troops were inadequately trained for mobile warfare, the margin of German qualitative superiority was probably greater on the Russian front than in France.

Employment doctrine, however, also played a role in the disaster that befell the Red Army in late June and early July. Driven by a doctrine of firm territorial defense and the offensive, Soviet commanders prepared a linear array directly on the USSR's borders. But with only 135 divisions along the lengthy border, and without major fortifications, the Red Army lacked the density to maintain an inflexible linear array against the kind of attack that

[14]See Natkiel, 1982, pp. 118–123.

the Germans could mount. In the sector north of the Prippet Marshes, the Germans massed enough forces to achieve an overall advantage of 1.25:1 in maneuver units; to the south, the Soviets enjoyed a similar edge but one they were not capable of exploiting. By concentrating further at three specific zones in the north, the Wehrmacht achieved local advantages of about 3:1 in maneuver units against a Soviet front line that, by doctrinal standards, was too thin to manage these disparities.

The frontal wall that the Red Army did construct consumed all but about ten divisions, thereby leaving almost no strategic reserve. Most of these divisions were deployed effectively around Minsk, but they were too small to block the German pincer movement directed at them. To compound its problems further, the Red Army deployed 30 of its tank divisions into the front line, leaving only two divisions in strategic reserve. Two-thirds of these 30 divisions were deployed south of the Prippet Marshes, leaving a 1:1 armor balance in the North. The Soviets dispersed the 12 armor divisions there among their five armies. The Germans concentrated their armor heavily. As a result, the Germans enjoyed armor advantages of between 3–5:1 in all three zones of their major attacks. This unbalanced distribution of armor, coupled with the Red Army's lack of mobile reserves, goes a long way toward explaining the major breakthroughs that the Wehrmacht was able to gain and then translate into sweeping maneuver victories in the rear areas.[15]

In summary, employment doctrine played a large role in the Red Army's defeat in Barbarossa. Given the lengthy border and its low ratio of force to space, the Red Army probably would have been better off to cede territory at the onset. An alternative would have been a mobile defense, with many more armored divisions held in deep reserve and employed en masse to check the Wehrmacht's advance. At a minimum, this might have bought the USSR additional time to mobilize and helped stop Hitler's drive well short of Moscow and Leningrad.

Whatever the case, linear defense would probably not have been a viable option during Barbarossa regardless of how the Russians distributed their forces along the front line. What separates Barbarossa from the Battle of France is that the allies had a sufficiently large army to contemplate a linear defense along the much shorter front and less trafficable terrain there. Overall, their force-to-space ratios were well within tolerable limits, especially taking into account the rivers, forests, and Maginot Line at their disposal. Yet the allies managed to lose by misallocating their forces along this linear defense and then failing to adjust their deployment pattern once it started breaking down. They lost less because of their linear array than because of its positional nature.

These two cases thus are similar in one respect, but different in another. Both involved the decisive defeat of a linear array at parity by a German attacker that showed mastery of such traditional principles of war as surprise, concentration, economy of force, unity of command, combined arms, and mobility. They differ in the sense that Barbarossa is a case illustrating how a defending army, fighting at parity, can fail because it had insufficient forces to contemplate a linear array in the first place. The Battle of France takes the point one step further by illustrating how an army, again at parity but this time amply sized for a linear defense, can lose by executing this concept poorly. It suggests that while adequate defensive size in relation to space might be a necessary condition for a linear array to work, it is not

[15]Natkiel, 1982.

sufficient. The capacity to act intelligently, to adopt flexibly, and to move quickly as the tide of battle changes is also important, indeed vital.

A MIXED OUTCOME: THE BATTLE OF THE BULGE

The final battle provides a case study in which the attacker penetrated the defender's front line and initially drove into the rear. But he was unable to exploit, largely because the defender mobilized sufficient reserves to restore a cohesive line and contain the attack. This case provides an opportunity to examine the interaction between force levels and employment doctrine from the vantage point of both success and failure.

The Battle of the Bulge was fought in December 1944. Although the three-year period between Barbarossa and this engagement saw a great deal of offensive warfare, few of the specific battles suit the purposes of this study: parity conflicts in which a breakthrough strategy was employed against a linear defense. On the Russian front, after stopping the Germans at Stalingrad, the Red Army went over to the offensive in late 1942 and began its long march toward Berlin. Over the next two years, it regularly employed concentration and breakthrough operations to breach German lines, drive deep into the rear, and compel the Wehrmacht to fall back from one river line to the next. But by this time, the Russians had amassed 450 divisions, giving them an overpowering 3.5:1 advantage over the Germans. In addition to the Russian front not being a parity fight, the Wehrmacht defense doctrine was hardly linear in the classical sense. The Germans tried to maintain a stable front line, but they executed it in a highly flexible way by relying on strongpoints and mobile reserves.

The war on the western front, meanwhile, was dominated by the broad-front offensive that Eisenhower pursued after the Normandy landing in June 1944. There were two important exceptions to this pattern. In late July, American forces broke out of the Normandy beachhead at St. Lo, executed a deep envelopment, and trapped a portion of the German 7th Army at Argentan. Then in September, British and American forces attempted Operation Market Garden, a daring but unsuccessful effort in the Netherlands aimed at turning the northern flank of the German West Wall. These operations aside, the allied drive to Germany's borders that summer and fall was achieved by broad front means in which allied forces steadily drove back Wehrmacht units that were trying to retreat to more favorable terrain.

By late fall 1944, allied forces had built to the point where Eisenhower now had some 65 combat divisions under his command. The German force opposing him originally had numbered only some 45 understrength divisions. After Hitler decided to launch a major counterattack in December, he refurbished these units with new troops and equipment. Additionally, he was able to mobilize the German economy to create 20 wholly new divisions. By early December the Germans had 65 full-strength divisions of their own on the western front. The new units were inexperienced, and many of their recruits were below normal Wehrmacht standards. Also, the allies enjoyed a major edge in air power as well as a larger manpower pool for replacing casualties. The overall balance thus was probably about 1.5:1 or more in the allies' favor. But at least on paper, the ground balance in maneuver units was about 1:1.[16]

Hitler's strategy was to concentrate large forces in the Ardennes, crash through American defenses there, and drive 160 kilometers to Antwerp. This maneuver was designed to cut

[16]See Eisenhower, 1986, pp. 554–661. See also Bradley, 1951.

off some 30 allied divisions deployed north of the Ardennes, among them 17 British and Canadian divisions. Senior Wehrmacht generals, doubting the feasibility of this plan, wanted to pursue a less ambitious effort that would penetrate to more shallow depths and entrap a single American army (6–9 divisions). But Hitler was thinking strategically, not tactically. His goal evidently was to drive Britain and Canada out of the war, thereby leaving the Americans isolated on the continent and facing the more experienced Wehrmacht alone. This political design, borne of Germany's increasingly desperate straits, left him favorably disposed to taking great risks.[17]

Brushing aside his generals' apprehensions, Hitler selected an area in the Ardennes just north of where the Wehrmacht had gained its breakthrough in 1940. His employment plan called for some 25 German divisions, including ten tank divisions, to be concentrated along a 110 kilometer sector between the Roer Dam and Luxembourg City. The assault was to be spearheaded by the powerful 6th Panzer Army, which deployed on this sector's northern flank. On its left, the 5th Panzer Army was to provide operational support by marching abreast. On the 5th Panzer's left, the 7th Army was to assume blocking positions to screen any American counterattack from the south. To the far north, German Army Group H was to pin down the American, British, and Canadian forces opposing it, leaving them vulnerable to entrapment from their rear.

The rugged terrain in the Ardennes favored the defense. Also the cold, snowy weather posed a further impediment to swift offensive operations. The Germans had plenty of experience with both the Ardennes and winter combat, something the allies lacked. Moreover, the prospect of an overcast sky promised to keep the allies' powerful air forces temporarily out of the sky, thereby allowing the Wehrmacht to operate free from air harassment. For these reasons the German high command judged the conditions to be favorable for an assault.

Assembling its forces in secrecy, the Wehrmacht hoped to gain surprise against a weak point on the allied line and then to capitalize on the allies' lack of reserves in the area. Eisenhower's posture was due to build to about 90 divisions by the following spring, which would give him a great deal more force in relation to the terrain being operated on. At the time, he was trying to attack with only 65 divisions along a front that extended from Rotterdam to Switzerland, 700 kilometers. The large-scale introduction of mobile equipment had enabled allied divisions to expand their frontages somewhat beyond the maximum length of 8–10 kilometers that had prevailed in 1940. But even so, Eisenhower's broad front strategy and linear array had compelled him to deploy nearly all his divisions on the front line, leaving only three divisions as a strategic reserve. These included two airborne divisions and one armored division that, only recently arrived, was still at the French docks unloading its equipment.

Eisenhower's linear array by no means was uniformly distributed, the result being a lack of forces in the area that the Wehrmacht planned to attack in strength. In order to marshall the local strength needed to push back the Germans along the front, the allies had concentrated in two sectors immediately adjoining the Ardennes. North of the Roer Dam, 20 allied divisions were gathered on a 75 kilometer sector. South of Luxembourg City, Patton's powerful 3d Army deployed nine divisions along an 80 kilometer front, supported by the 7th Army's nine divisions on a second 80 kilometer front immediately below Patton. But in the Ardennes the Americans had only the equivalent of six divisions from Hodges' 1st Army

[17]MacDonald, 1985.

guarding the 110 kilometer sector there. Table 14 displays theater deployments for both sides.[18]

The Battle of the Bulge was heavily influenced by the manner in which this allied theater deployment interacted with the Wehrmacht's strategy of concentrated attack. This interaction produced a battle in two parts. The first part was the forward battle, which the Germans largely (but not entirely) won because they overpowered an outnumbered U.S. opponent that did not fight well in places. The second part was the rear area battle. It was won by an allied army that drew on its large forces fortunately deployed on both sides of the Ardennes to counter-concentrate faster and more effectively than the Germans thought possible. The outcome was the "bulge" for which this battle was named: a triangular salient that grew to a depth of 90 kilometers over the first week, then stalemated as the German attack lost steam, and finally began contracting as the allies counterattacked.

When the battle opened on the morning of December 16, the American contingent in the Ardennes found itself facing a 25-division attacking force that outnumbered it by about 4:1 in maneuver units and manpower and by nearly 7:1 in tanks. The Americans benefitted from favorable terrain features, including the Schnee Eifel, Skyline Drive, and the Our, Clerve, and Sure rivers. Consequently, they were not spread hopelessly thin. But their average divisional frontage of about 18 kilometers did exceed doctrinal standards by a substantial margin. As a result, virtually all battalions were committed on line, leaving almost no local reserves. Compounding their problems further was the fact that the two divisions and one cavalry group on the northern shoulder were clustered on a 25 kilometer front. This left the divisions to the south defending frontages averaging 25 kilometers. Table 15 displays frontages for the U.S. Army units deployed in the Ardennes on December 16.[19]

Most of these divisions were not well-prepared for battle. One had only recently arrived in Europe and had not yet seen much combat. The others were combat-trained but had recently been exhausted by bloody fighting in the Hurtgen forest and had been sent to the Ardennes to recuperate. They were in the process of absorbing many replacements, all raw recruits. All these factors added up to a vulnerable situation in the face of a determined German attack plan to create the quick breakthrough that Hitler's strategy depended on.

Table 14

Allied and German Deployments on Eve of the Battle of the Bulge (divisions)

	Length (km)	Allies	Germans
Far north	150	6	7
Immediate north	150	24	8
Breakthrough sector	110	6	25
Immediate south	110	13	11
Far south	200	13	9
Strategic reserves	—	3	5
Totals	720	65	65

[18]Eisenhower, 1986, p. 545.

[19]MacDonald, 1985, pp. 101–196. These frontages are approximate. Other historical accounts may show somewhat different frontages from one division to the next, but the basic pattern remains constant.

Table 15

U.S. Army Frontages in the Ardennes, December 16, 1944

Unit	Frontage (km)
102d Cavalry Group	5
99th Infantry Division w/ 2d Infantry Division in reserve	20
14th Cavalry Group	7
106th Infantry Division	20
28th Infantry Division	31
9th Armored Division (2/9)	3
4th Infantry Division	24[a]

[a]The remainder of the 9th Armored Division was deployed to the north of the Ardennes sector. Thus six "division flags" were present at the Ardennes; counting the two cavalry groups, the total force was about six DEs.

Although the Germans successfully gained surprise, their initial progress on the first day was not uniformly encouraging. On the critical northern shoulder, the 6th Panzer Army ran into greater opposition than had been expected. Barring its way were two commingled American divisions (99th and 2d), rather than the one division that Wehrmacht intelligence had mistakenly detected. In the center, however, the Germans got the opportunity they were looking for. In the Losheim Gap, the 14th Cavalry Group, lightly armed and dispersed among several villages, quickly withdrew under the weight of a German artillery, infantry, and armor assault. On the 14th Cavalry Group's right, the 106th Infantry Division soon found both its flanks penetrated. When a command foul-up led to a failure to withdraw the division, two of its regiments were encircled and bypassed. Two days later, they surrendered. Further south still, the thinly spread 28th Infantry Division fought hard and initially held the Germans at bay. Badly weakened and running out of ammunition, it finally buckled at Clervaux on December 18.[20]

The defeat of these units opened up a gap 40 kilometers wide that enabled several German armored columns to pour into the rear areas along three road arteries and to fan out. The Ardennes battlefield erupted as German and American forces, both trying to assert control over the battle, engaged in a fierce contest marked by sharp fighting, continual movement, and improvised tactics. Bypassing and outflanking American positions along the way, Wehrmacht units tried to advance as rapidly as possible, sowing confusion into U.S. Army forces along the way. American tactics were aimed at restoring cohesion, holding critical communication nodes and arteries, and reestablishing a contiguous front line. These efforts eventually were to succeed, but initially they were disorganized, uncoordinated, and scattered randomly across the swirling battlefield's landscape.

American units, often widely separated from each other, fell rapidly backward as key towns and road junctures were ceded. By December 20, the German main effort had advanced about 35 kilometers. American forces had fallen back into pockets at St. Vith and

[20]MacDonald, 1985, pp. 466–487.

Bastogne, separated by a 20 kilometer gap, and the Bastogne salient was not linked to American defenses on its southern flank. Since both pockets sat astride important road junctures, they were to play a vital role in delaying the German advance. On December 21, nonetheless, two German armored columns swung around Bastogne on each side and began driving toward Dinant. That same day the St. Vith pocket crumbled and American forces there retreated northward, enlarging the corridor to the west. With the Germans steadily approaching the Meuse River, the allied position began to look increasingly precarious.

Over the next week, however, the rear area battle turned in the allies' favor and the situation stabilized. The German drive first slowed and then came to a complete standstill. At their furthest point the Germans on December 26 came within 10 kilometers of the Meuse. But by this time their salient, pressured by allied forces on their northern flank, had narrowed to the point where a single British armored brigade blocked it outside Dinant. Further to the north, where the German drive originally was to have been channelled, Wehrmacht units fell 40 kilometers short of the Meuse. To the south, meanwhile, Bastogne was liberated and American forces there rescued.

The German breakthrough initially had all the earmarks of 1940 including an audacious thrust, covering 30 kilometers in one day, by a single S.S. armored brigade under the daring Col. Joachim Peiper. What happened to stop it? Many factors were at work, including some on the German side. The Germans were constrained by logistic supply problems and traffic snarls that dogged them as the battle wore on. Also preventing the Germans from repeating their 1940 masterpiece were less brilliant commanders, tactical errors, untrained units, the absence of air support, and the lack of enough combat forces. For the allies, a particularly important factor was the stubborn way American forces fought in retreat after their frontal defenses had been penetrated. While American forces there did not win, neither did they allow the battle to deteriorate into a rout.

Despite the unravelling of the American defense line from December 16 onward, most U.S. Army units caught within the Bulge maintained their composure. Unlike the French Army in 1940, nearly all were well-armed, especially with antitank weapons that enabled the infantry to stave off repeated armor attacks. And they had more tanks at their disposal, which fought with greater skill than the French tanks had. Even though often isolated and outnumbered, they kept fighting, and, unlike their French counterparts, to a great effect.

The Germans particularly ran into trouble on the Bulge's northern shoulder, where American defenses remained strong from the first day. German strategy hinged on the 5th Panzer Army's ability to overpower American defenses there and then pivot northward toward Antwerp. The American units were strong enough for a stalwart defense and quickly reinforced, blocking the German main effort and channeling it southward, away from Antwerp. Especially important was the Battle of Elsenborn Ridge on December 19–21, where savage American artillery fire shattered four attacking German divisions. To the west and south, isolated American units also threw sand in the gears of the German advance. A few engineer units frustrated Peiper's brigade by blowing up several bridges in its way, ultimately leaving it exposed to destruction by American reinforcements. Further south still, the isolated American pockets of resistance at St. Vith, Bastogne, and elsewhere continually blocked and disrupted the Germans.[21]

These efforts helped give allied commanders time to mobilize and redeploy large reserves, the dominant factor behind the allies' success in the rear area battle. In this respect, the allied response also differed from that of the French in 1940. The French had

[21]MacDonald, 1985. Also, see Bradley's (1951) description of allied defense tactics during this battle.

wholly failed to bring large reserves to bear in a timely fashion; the allies succeeded. In important ways, the battle turned on this response. Had the French done better at counter-concentrating reserves, the 1940 conflict might have turned out more favorably. And had the allies done less well, the Battle of the Bulge might have ended tragically.

Initially the same kind of paralysis seemed to be settling into the allied command structure that afflicted the French in 1940. Before the batttle's onset, allied leaders were not completely lacking intelligence indicators and advanced warning about the German buildup in the Ardennes. Despite the tentative concerns of a few, however, senior allied commanders evidently failed to anticipate that the Germans, who seemed nearly beaten, would try anything so bold. Once the battle began, several hours passed before the realization began sinking in that a sizable Ardennes offensive was underway. Communications were disrupted by the German artillery bombardment and time was required before enough information could filter back to headquarters staffs to provide a coherent picture of what was transpiring.

Shortly thereafter local corps commanders began issuing orders to alert their forces and move units to the general scene. They also began calling for reinforcements from outside the Ardennes sector. Senior commanders (Bradley and Eisenhower) honored these requests and began crafting a defensive battle plan aimed at building a line on the Germans' northern flank and establishing strongpoints to block them from moving west. Nonetheless, they initially regarded the attack as a local spoiling action. Skepticism about German capabilities was one factor in their judgment. Another was lack of clear evidence that the Germans had yet decided to commit a large force. At the time, only a few German divisions had entered the battle; the rest remained in rear areas from where they might quickly be shifted elsewhere. A third factor was reluctance to disrupt allied operations actions elsewhere in the theater, especially the offensive actions near the Roer Dam and in Alsace. It was not until December 18th, two days later, that they decided a major offensive was underway, and it was the 19th before they began taking command decisions to reorient the entire theaterwide posture.

In other circumstances, these delays might have been fatal, but the decisions taken at lower levels in the first day or two did result in the prompt shift of sizable reinforcements toward the Ardennes. By late on the 19th, less than four days after the battle started, American strength in the Ardennes already had doubled. What began as a six-division posture now stood at 13 divisions. Allied armor had grown from one to four armored divisions, enough to reduce the imbalance from 7:1 to only 2.5:1. Four of these divisions came from Hodges' 1st U.S. Army, which was deployed north on the Roer Dam. Two others were the 82d and 101st airborne divisions, then in theater reserve status. The final division was drawn from Patton's left flank, immediately south of Luxembourg City.[22]

On the vital northern shoulder, two extra forces helped shore up the allied position from the Elsenborn Ridge westward. At St. Vith, the 7th Armored Division and the 82d Airborne anchored the pocket that delayed the German advance three critical days. At Bastogne, the 101st Airborne enabled the Americans to hang onto the important road juncture there. Finally, Patton's 10th armored division helped shore up the southern flank around Luxembourg City.

British and American reserves provided a further infusion. On December 19th, the senior commanders met to devise a theater strategy. The past two days had seen the commitment of more German divisions and their aggressive march westward, which removed

[22]Eisenhower, 1985, pp. 554–661.

lingering ambiguity about Hitler's intentions. Having concluded that the Germans were pursuing a major drive, Eisenhower ordered a halt to all allied offensives in Europe to focus on the Ardennes. Despite their concern about the reversals over the past four days, the group felt confident about the allied ability to restore the situation. The issue was exactly how to react.

Patton wanted to launch quick allied counterattacks into the German flanks, but his colleagues demurred. Among their reasons was concern that early counterattacks might be premature and uncoordinated, thereby leading to piecemeal commitments that might be chewed up sequentially. The group reached a consensus in favor of a more conservative strategy aimed at shoring up the allied line around the expanding salient especially in the north, where the German advance posed the greatest danger. In the south they agreed upon an early but limited counterattack to relieve Bastogne. The goal of this strategy was to contain the German advance and channel it away from vital areas. Once the German attack had lost energy, the group agreed, a full counterattack aimed at destroying the German forces and recapturing lost ground would then be possible.

To implement this strategy, Eisenhower ordered a deployment of additional allied forces to the Ardennes. From the south Patton was to detach a three-division corps from his 3d Army, turn it northward, and drive to Bastogne. Devers' 6th U.S. Army Group was to shift its boundary northward to cover the gap created in Patton's line. From the north, Simpkins' 9th U.S. Army was to detach two divisions and move them southward, while also covering a portion of Hodges' line to give the 1st Army greater scope for maneuver in the Ardennes. The U.S. division at the dockyards in France was to speed forward to the Bulge. Finally, Montgomery was to move the 30th Corps of the British 21st Army Group, with four divisions, into position on Hodges' right flank to block German access to the Meuse. Other British divisions were to form further to the rear, providing a second screen.[23]

Taken together, these additional commitments brought the allied posture slotted for the Ardennes to a total of 25 divisions. Four other divisions were held in tactical reserve. The 25 committed divisions were enough to achieve a 1:1 ratio in maneuver units and to establish a cohesive, reasonably dense line along the salient's full length on both sides. Ten of these units were armored divisions, thereby bringing the local tank balance into equality or better. These commitments placed the allies in a position to execute their defensive strategy and thus bring the German advance under control.

The employment scheme by which this strategy was implemented is displayed in Fig. 35, which shows how U.S. and allied divisions were lined up along the Bulge on December 24. By reestablishing a linear array on an east-west axis, this employment pattern blocked a deep German advance in all directions, while remaining elastic enough to give the Germans room to dissipate their energy harmlessly. As events turned out, this employment fulfilled the requirements of Eisenhower's strategy and ultimately gave the allies victory in the defensive stage of the rear area battle.

With the German drive mired at Dinant and clear weather now permitting air operations, on January 3 the allies embarked upon their own offensive stage by launching a full-scale counterattack. In a coordinated fashion, allied forces on both sides of the Bulge began pushing toward the center and eastward. Faced with this counterattack, Hitler finally acknowledged that the Ardennes offensive had failed. On January 8, the Wehrmacht, now facing entrapment itself, began withdrawing. By January 28, the allied line had been restored to its original position, sending the Battle of the Bulge into history.

---

[23]Eisenhower, 1985.

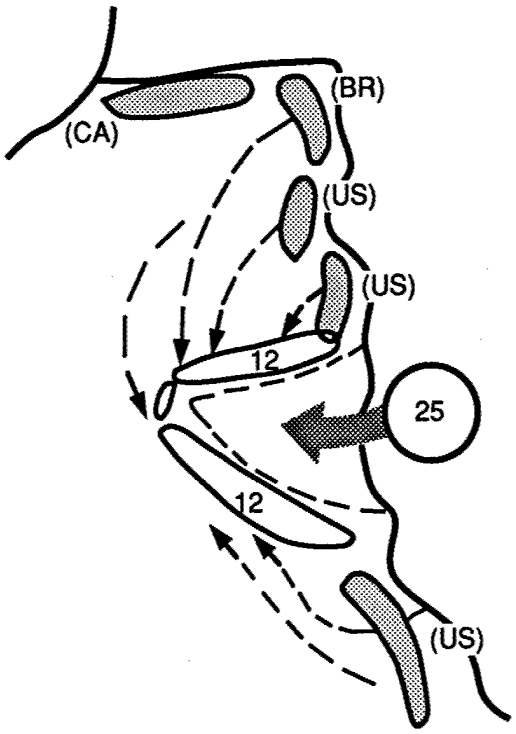

Fig. 35—Allied Defensive Employment Strategy (Divisions on December 24)

Over the entire battle, the allies suffered 83,000 casualties in killed, wounded, and captured. The Germans lost about 100,000 in all three categories. The exchange ratio thus was only about 1.20:1 in the allies' favor, well below the 2–3:1 ratio expected in the theory of linear defense. In operational terms, however, the Germans suffered heavily. While the allies were able to replace their losses within a few weeks, the Germans lacked comparable ability to replenish either equipment or manpower. The weakened Wehrmacht was unable to put up stout resistance when the allies resumed the offensive in February. Quickly breaching the west wall, the allies crossed the Rhine on March 7 and by April 1 had captured the industrial Ruhr. When the war ended on May 8, allied armies stood on the Elbe River, 100 kilometers from Berlin.

What is noteworthy about the allied defensive strategy at the Battle of the Bulge is not only its basic design, but also the fact that the allies were able to implement it so quickly. On December 16, the allied posture in the Ardennes stood at only six divisions. By December 19, it totalled 13 divisions. By December 24, only five days later, nearly all 25 divisions were in place. On December 26, the ring around the Bulge had been firmly set. Part of this success story owes to the tactical mobility of the allied forces, which impressed German officers. More heavily mechanized and motorized than allied forces in 1940, British and American units in 1944 simply had faster cross-country movement rates, allowing them to redeploy faster. Particularly remarkable was the performance of Patton's 3d Army, elements of which moved from Alsace to the Ardennes in only five days. Ordered to redeploy on December 19,

Patton's forces were ready to move by December 21. Turning northward, they marched some 250 kilometers in less than three days, hit German lines on December 24, and liberated Bastogne 48 hours later.[24]

Patton's dash north involved three divisions, amounting to only about 15 percent of the allied forces that converged on the Ardennes battle over the span of a few days. The remainder got there so fast less because of their tactical mobility than because they had only a short distance to travel. This was particularly true for the ten divisions from the U.S. 1st and 9th armies, several of which had arrived by December 19. Originally concentrated directly on the battle's northern flank, they had to travel only 50–75 kilometers to get there. Also, Montgomery's British forces, deployed immediately north of the U.S. 9th Army, were themselves only about 100 kilometers away. Under these conditions, a rapid allied counter-concentration became possible.

An additional factor helping the allies is that these reserves were physically able to disengage themselves, with little or no delay, from battle in their own sectors. When the Ardennes offensive began, the 1st and 9th armies were so heavily concentrated for the impending offensive around the Roer Dam that they were echeloned in depth. Two thirds of their 15 divisions already were positioned in the rear and therefore did not have to disengage from frontline assignments. Their ability to turn immediately southward accelerated their arrival by a day or two, a difference that might well have been critical. Montgomery's and Patton's forces both were committed on line at the time, but neither were under pressure from the Germans in their sectors. They too were able to disengage quickly, turn toward the Ardennes, and move free from harassment by enemy air strikes. Greater German ground and air pressure might well have made their response more sluggish. Finally, the allied reserves were able to converge on the Ardennes from three directions. Patton moved north, Hodges directly south, and Montgomery in a southwesterly direction. As a result, they did not overload available roads in ways that otherwise might have slowed their response.

The allies thus benefitted from a stroke of good fortune in the nearby proximity of large reserves that physically were free to redeploy. Different circumstances could have produced a less effective result. For example, had a larger contingent been deployed further north, which Montgomery had wanted, considerably fewer would have been promptly able to reach the Ardennes quickly. Eisenhower's broad-front strategy, its other liabilities aside, served the allies well at the Battle of the Bulge.

SUMMARY OBSERVATIONS

Perhaps the most fundamental insight these historical cases reveal about the interaction between linear defense and offensive strategies of concentration, breakthrough, and maneuver is that attacking armies have often resorted to a maneuver strategy when facing a linear defense. Even in situations of parity, they regularly have done so in the expectation that a decisive victory was at least possible. The very fact that their efforts succeeded sometimes but failed at others, however, suggests that the outcome in any single case is not foreordained.

The theory of first fracturing the defender's cohesion to wage a climactic battle in the rear has a strong empirical and historical base. More than an abstract theory, it often has enabled an attacker to destroy a defender while suffering only light casualties in return. At

[24]Natkeil, 1982, p. 177.

the same time, World War I indicates that the theory of linear defense is feasible when the conditions are right. A maneuver strategy requires a breakthrough. When a breakthrough cannot be gained because the defender's frontal wall is too strong, this strategy is doomed to failure.

World War I was unique; front lines are seldom this strong and are equally unlikely to be unbreachable in the future. The intriguing question is what happens when a local breakthrough does occur, and why? The Battle of France and Barbarossa show that the outcome can be a disaster for the defense. But the Battle of the Bulge shows that if the defender can keep his balance, he can still salvage the situation and even emerge with a victory of his own. What determines this wide range of variation?

There is no simple, single-cause answer to this question. A quick comparison of the Battle of France and the Battle of the Bulge does at least help identify the multiple variables that seem important. These two battles were fought on similar terrain and almost in the same place. In both cases, a powerful, armor-heavy force attacked on a 100 kilometer sector against a linear array and quickly gained a breakthrough. In the former case, the attacker rolled to victory; in the latter case, he was frustrated.

At first glance, the attacker's own dissimilar strength levels seem to have played a big role in determining why these battles ended differently. The Germans committed 45 divisions to the Ardennes breakthrough in 1940, a larger force than the 25 divisions employed in 1944. They had a more powerful force available in 1940 and a greater local margin of superiority over the defender. Aware of this, the German high command evidently believed, before the 1944 attack began, that another ten divisions were needed to provide a good shot at success.

Weighing against this factor as the dominant explanation are that the Germans did not use all their forces in 1940 and their 1944 forces did not fail primarily because they were too small. In each case, the breakthroughs were caused and initially exploited by a lead echelon numbering only 10–15 divisions. The German 2–3:1 edge at the main point of contact proved adequate to gain quick penetrations both times. Additionally, the 1940 force was no bigger than the 1944 contingent in relation to the size of the enemy force ultimately to be engaged. In both cases, the Germans sized their attacking and pin-down forces together to achieve about a 1.25:1 advantage against the portion of the defender's army that was to be entrapped. In this relative sense, the two attacking armies were very similar.

The difference cannot be ascribed wholly to the less impressive performance of the German Army in 1944. In 1940, the cream of the German Army attacked the French. But historical lore notwithstanding, the Germans did not perform perfectly even then. For example, their tanks did not match well against allied tanks and they made some tactical errors on the road to Dunkirk. By 1944 the Wehrmacht knew a great deal more about armored warfare and it had superior weapons, including better tanks. The German Army of 1944 might have fought with less élan, but its Ardennes offensive showed no less daring. Nor was it pursued incompetently. Once the breakthrough had been achieved, German armored columns advanced with their customary vigor. In 1940, the Germans stormed 60 kilometers in three days; in 1944, they gained about 35 kilometers in four days. The difference between the two was important, but it was brought about primarily by greater allied resistance in 1944 rather than by any German deficiencies.

Where the German attack in 1944 suffered by comparison was in overall theater employment strategy. In 1940 a German force of 30 divisions marched forward on the right flank of the Ardennes force. Its task was to lure French and British forces into Belgium and

then to pin down 60 allied divisions on the Dyle River, well north of Sedan. It accomplished this task remarkably well, performing an act of military jujitsu that was abetted by the allies' own willingness to swing their left flank forward. In 1944 a German force of 15–20 divisions stood on the right flank of the Ardennes contingent. Facing 30 allied divisions, it was no worse off in relation to the enemy than its 1940 counterpart, but it failed to have a comparable effect.

The German high command originally did plan a supporting attack north of the Ardennes. Shortly after the battle got underway, however, Hitler cancelled it to divert his five strategic reserve divisions to the Ardennes. Even had this secondary attack been launched, major barriers would have stood in its way. British and Canadian forces westward of Nijmegen, the corner of the allied line, were spread thinly with only six divisions guarding a 150-kilometer front. But they also were screened by the Maas and Waal rivers, which inhibited a southward thrust from there. The only other alternative was the area between the Roer Dam and Nijmegen. That was the sector the allies themselves had chosen for their own concentrated offensive, and the presence of 24 allied divisions there inhibited a German assault. As a result, the Germans attacked in the Ardennes with little pin-down pressure on their right flank and with virtually no prospect that allied forces there could be lured forward.[25]

Even so, the Germans attacked in the Ardennes itself with great intensity, and the fighting was touch and go for several days. The primary difference between 1940 and 1944 lies in the far greater quality of the defense that the allies were able to mount after the initial breakthrough.

The difference, to be sure, was not evident in the first day or two. The frontally committed forces of both the French and American armies buckled quickly under the Wehrmacht's blows. Neither force, on paper, should have crumpled so fast. Although thinly deployed, both were far from being porous and therefore had enough raw combat power to hold up longer than one or two days. Modern divisions, when deployed on manageable frontages, are typically sized to be able to fight effectively for three to seven days, and maybe longer, even against a strong attacker. But both forces gave way and for reasons that often had little to do with paper strength. One French division was destroyed in a violent attrition battle that did reflect material strength, and two others simply unravelled as their troops panicked and fled. As for the Americans, one division also succumbed to attrition, but a cavalry unit was driven out of position while its neighboring division was removed by being flanked and encircled. Both battles demonstrate that especially when tactical reserves are not available, breakthroughs can be caused by a good deal more than attrition dynamics and that they can happen much faster than expected.

There, however, the similarity between the two defense efforts ends. The French units between Namur and Sedan in 1940 seem to have disappeared after they had been penetrated. Having lost their unit cohesion, they no longer acted as a serious impediment to the Germans. The American units, by contrast, remained a factor in the battle. Although driven backward, they continued to fight effectively and thereby bought valuable time for the defense to form further back. Moreover, a coherent defense did take shape in the rear areas because sizable reserves promptly converged on the scene. The difference between the two battles is illuminated by Table 16, which displays the buildup rate of French and American reserves in these two battles, a powerful explanation for the fact that the Germans' 1944

[25]MacDonald, 1985, pp. 604–619.

Table 16

Allied Reserve Buildup Rates (cumulative total in divisions)

	Battle of France	Battle of the Bulge
D-Day	0	0
D+3	2	7
D+5	3	14
D+8	6	20

drive was stopped only 90 kilometers into allied territory, whereas their 1940 drive carried them to the English Channel.

The Battle of the Bulge thus offers a prescriptive model for how a linear defense can absorb and contain a concentrated attack that initially does gain a breakthrough but that successful linear defense is no easy task. As in 1940, the nature of the German attack came as a surprise, the breakthrough was gained with blinding speed, and the battle exploded into the rear areas with a stunning ferocity. The American forces in the Ardennes held on by their fingernails, the timelines were narrow, and for several days the outcome was in doubt. The allies ultimately prevailed because they redeployed large forces quickly. Their ability to do so, however, stemmed not only from their own agility but also from the unique features of the battlefield and their own theater employment configuration. As Table 17 shows, the allied posture was not deployed in a perfect linear array that would have conformed to the literal dictates of a broad-front strategy. Large allied forces already were concentrated on the shoulders of the Ardennes, thereby creating a pool of some 37 divisions available for quick redeployment. A real linear array would have reduced this pool to some 24–26 divisions, a big difference. Had these extra forces not been located nearby, and had the allied far northern flank not been screened by multiple rivers, the allied response would probably have been less effective. At a minimum, the allies would have had to mount an even more complicated and demanding redeployment effort.

The Battle of the Bulge's eventual outcome should not be taken as an iron-clad prediction of how tactically mobile defense forces will perform in the future. Although the skill of the defender's forces in the breakthrough zone is important, the defender's overall pattern of theater employments matters as well. Concentration near the main point of contact begets an agile response; the converse is also true. The battle further suggests that while the defending army might not have to be combat-trained, previous experience matters. Finally, once a breakthrough has occurred, more than a few reserve units may be needed to restore control. The allies promptly assembled enough reserves to reconstitute a defense line along a deep salient and to achieve a local 1:1 ratio of forces. Without better information on the subject, this might not be a bad standard for applying to the defense's force requirements of the future.

Table 17

Allied Force Deployments vs. a Real Linear Array (divisions)

	Sector Length (km)	Actual Number of Divisions	Linear Array with Weak Ardennes	Linear Array with Uniform Doctrine Densities
Far north	150	6	13.8	12.9
Immediate north	150	24	13.8	12.9
Breakthrough sector	110	6	6.0	9.5
Immediate south	110	13	10.1	9.5
Far south	200	13	18.3	17.2
Total forces		62	62.0	62.0
Ardennes plus two adjacent sectors		43	29.9	31.9
Two adjacent sectors alone		37	23.9	22.4

5. LINEAR DEFENSE AT LOW FORCE LEVELS: A FORMAL INQUIRY

This section presents a formal, theoretical, and quantitative analysis of why and to what degree a NATO linear defense at low force levels would be vulnerable to enemy breakthrough attacks even in a situation of parity, or something like it. It is composed of three parts:

- First, it constructs a general conceptual model of breakthrough and maneuver operations against a NATO linear defense conducted at low force levels.
- Second, it conducts a static and dynamic analysis of how the prospects for a NATO linear defense change as a function of NATO's ground force size. It focuses on how NATO's prospects for a linear defense seem to worsen as its "force-to-space" ratio, or the size of NATO's forces in relation to the terrain being covered, declines.
- Third, it conducts a sensitivity analysis to determine how other variables affect linear defense's prospects, including the physical geometry of the battlefield, relative reinforcement rates, airpower and operational fires, command and control, and logistics support.

In the years ahead, NATO will continue to face the problem of defending lengthy borders that, unless the linear defense philosophy is changed, could spread NATO's forces dangerously thin. As Fig. 36 shows, NATO's defense plans will switch from defending the old inter-German border to protecting Germany's eastern borders on the new Oder-Neisse line. This change is a major one in political but not geographic terms. The new NATO defense line will be about as long as the old line, making it difficult for any NATO defense effort to simultaneously protect the entire border at uniform strength and respond to locally concentrated enemy attacks.

As Sec. 2 pointed out, NATO will no longer face the prospect of defending outnumbered by 2:1 as was the case during the Cold War, but it will still be outnumbered by some margin. Assuming that the Soviets remain capable of fielding an attacking army of about 50–65 divisions and NATO will retain a defending force of 40–45 divisions, NATO will face a numerical disparity of 1.25–1.50:1. The dynamic wargaming analysis presented in Sec. 2 concluded that NATO realistically could aim to defend successfully in the face of this disadvantage, but only if it abandons its linear philosophy and adopts an "echeloned defense." The advantage of an echeloned defense is that it considerbly enhances NATO's ability to move reserve forces to enemy breakthrough sectors in a timely fashion, thereby denying the enemy the capacity to wage a war of maneuver in NATO's rear areas. By preventing the enemy from pursuing a maneuver strategy, NATO would be able to confine the battle to the forward areas where it could benefit from the classical advantages that enable an outnumbered force to defend successfully.

An echeloned NATO defense therefore will be necessary if the future European force balance evolves along the lines estimated here. The matter, however, does not end there. NATO has committed itself to pursuing a second phase of arms control negotiations as a follow-on to the first treaty on Continental Forces in Europe (CFE I). Conceivably NATO

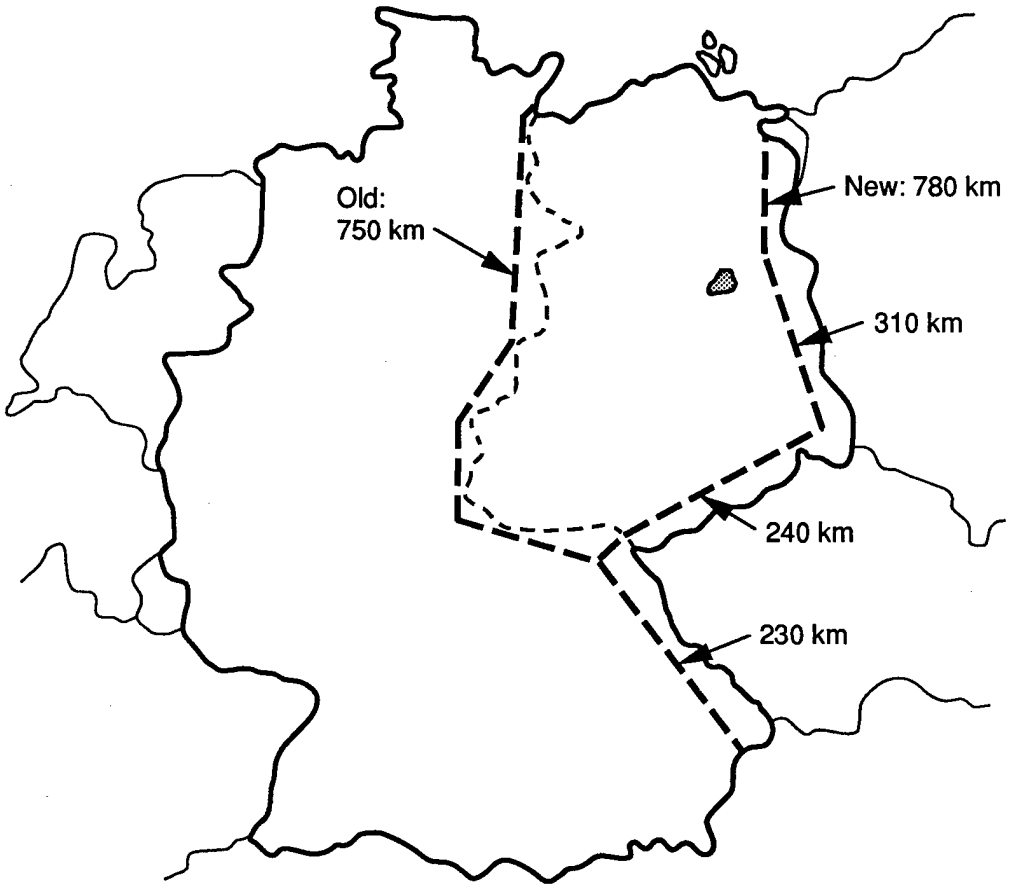

Fig. 36—Effect of German Unification on NATO's Military Geography

could use a CFE II negotiation to rectify the remaining imbalance through further asymmetrical cutbacks, producing a situation of genuine parity in ground combat forces. The process of negotiating such an accord would probably result in NATO reductions below the 40–45 divisions estimated here, further diminishing NATO's force-to-space ratio.

Would genuine battlefield parity bring about the rebirth of a NATO linear defense? Or would NATO's declining force-to-space ratio make linear defense still unwise even though the force imbalance has been rectified and parity achieved?

The historical analysis presented in Sec. 4 concluded that even at parity modern armies often have fallen victim to an enemy strategy of breakthrough and maneuver. This was the case when they attempted to implement a linear defense with a low force-to-space ratio and with a force posture and doctrine that did not allow them to respond in a timely fashion. The analysis presented here suggests that NATO could create similar problems for itself if it were to attempt to preserve a linear defense in Central Europe as its own forces are reduced. This will remain the case even if parity is achieved and the forces on both sides are reduced well below estimated levels (e.g., by 50 percent). The issues, however, are very complex, and deserve careful evaluation.

This negative conclusion about linear defense's future does not mean that NATO should refrain from future, post-CFE I arms accords that further reduce forces in Central Europe. NATO is politically committed to mounting a "forward defense" that endeavors to protect the borders of its members, including Germany, but a forward defense can be mounted in other than linear ways. One of these options is an echeloned defense that provides a combination of frontal strengths in critical areas and large, responsive reserves. A central implication of this section is that, whereas a rigidly linear NATO employment doctrine may be incompatible with a "deep cuts" arms control regime, a doctrine of echeloned defense can enable NATO to defend in the flexible, adaptive ways that are needed at low force levels.

An echeloned defense can work in tandem with a "deep cuts" arms control regime to enhance overall political and military stability in Central Europe. A cautionary word, however, is in order. Echeloned defense is not an easy doctrine for a multinational coalition alliance to master. And it carries risks, possible political sacrifice, and potential instabilities of its own. Careful NATO planning will need to address these problems, regardless of whether the Alliance pursues deeper force cuts in future arms control negotiations.

A GENERAL MODEL OF LINEAR DEFENSE AGAINST BREAKTHROUGH ATTACKS

Over the years, several studies have suggested that a properly designed NATO linear defense should be able to fight outnumbered by a margin of 1.50 to 1.75:1 and still rebuff an attack.[1] Why then, should NATO worry about its posture if the future Soviet/NATO force ratio will fall below this threshold? The answer is that this threshold is a variable, not a constant. It depends heavily on how large a force the defender deploys in absolute terms, and on whether that force is robustly endowed in relation to the missions that have to be performed in conducting a linear defense. As the defender's strength rises, the threshold increases; as his strength declines, the threshold decreases. The threshold of a 1.5–1.75:1 ratio applies to a large Cold War NATO posture of 50–60 divisions, which would be sizable enough to form both a strong front wall and ample reserves. At lower NATO force levels, this threshold declines; indeed, as Fig. 37 suggests, it converges on 1:1. In other words, force reductions strip away a linear defense's margin of safety and leave it lacking confidence in its ability to fight outnumbered. For that matter, parity itself would be no guarantee of success.

The general model to be developed here examines how a linear defense's prospects change as a function of different levels of defender strength. It postulates that modern ground warfare is affected not only by the force ratio (the relative strengths of the contending armies), but also by the force-to-space ratio. At one extreme is the situation that occurred in France during World War I. There, both the German and allied armies were so large that they physically dominated the battlefield. With over 150 divisions apiece, both armies were able to cover their entire front lines in a very high density (roughly one division per five kilometers) while still withholding sizable reserves. At the other extreme is the situation that prevailed in the early 19th century, when wars were fought by armies of only 100,000–200,000 soldiers and typically operated on wide open battlefields. In between these

[1]The analysis developed here primarily examines ground forces, with air forces relegated to the background. The underlying assumption for the air balance is that the two sides will retain equally large air forces that will engage in a competitive struggle while the ground battle is being waged. The analysis gives NATO credit for having a qualitative edge, especially in ground attack missions, but it also assumes a strong Soviet air defense effort that prevents NATO's air forces from asserting the kind of mastery over the battlefield that occurred in Desert Storm. Should NATO be able to achieve equivalent air supremacy, the risks described here would be less pronounced.

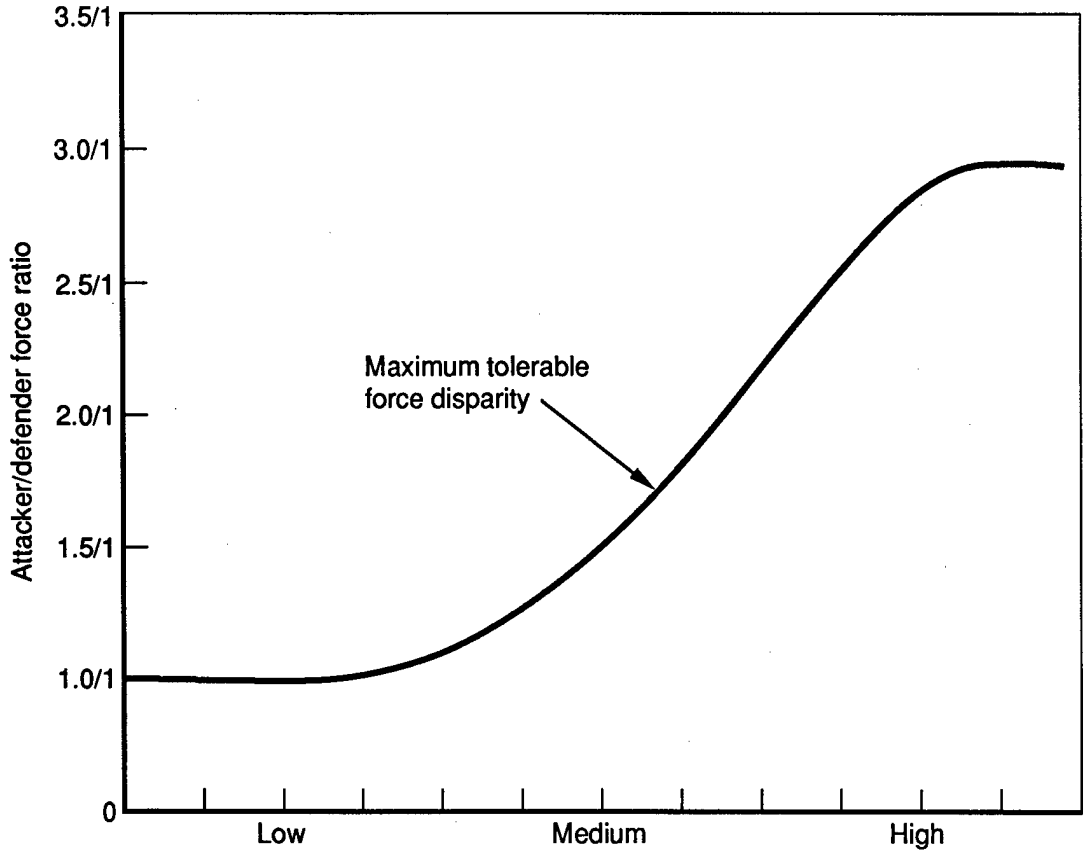

Fig. 37—Relationship Between Defender Force Levels and Defensible Thresholds

two extremes lies a wide spectrum of force-to-space relationships, including the present and future Central Region military balance.[2]

In general, the dynamics of ground warfare change as a function of altering conditions along this spectrum. In situations of high force density, there is little scope for operational maneuver. The attacking army is presented few geographic opportunities for moving its forces in ways to create decisive local advantages, or avoiding the defender's main strengths, or otherwise fostering climactic battles under favorable conditions. Warfare tends to become a grueling attrition process in which the outcome primarily is determined by whether one side can outlast the other. World War I evolved into this kind of classic attrition struggle. Because it was fought by two fairly equal, well-entrenched armies, it dragged on four years until the German Army finally collapsed from within. During its last two years, the Korean

[2]For this reason, RSAS models have been designed to include recognition of both this force-to-space dependence and the potential for explosive breakthroughs if the defender's forces are spread too thin.

War evolved into a similar attrition campaign with the United Nations and Chinese armies both dug into formidable mountain ranges, facing each other across the demilitarized zone north of Seoul.

In situations of low force density, by contrast, the scope for operational maneuver increases considerably, especially on the open, rolling terrain of Central Europe. With a portion of the battlefield terrain left uncovered by the defender, the attacker can aspire to gain not only a local advantage by concentrating its forces but also an open avenue for driving deep into the defender's rear areas. Once there, the attacker can pivot and advance laterally to separate the defender's units from each other, rendering them exposed to a sequential defeat in detail. The defender can also employ maneuver tactics of his own, tactics aimed not only at containing an attack but also at decisively defeating enemy forces. In this form of warfare, attrition remains the final arbiter, but maneuver establishes the all-important conditions of battle, conceivably paving the way toward a quick and decisive victory.

In situations of low force density, maneuver operations provide a larger attacker with a ready vehicle for overpowering a smaller defender. Equally important, in theory they also provide an attacker a militarily exploitable option for defeating an equal size, or even larger, defender, and vice versa. This especially is the case when one army enjoys qualitative advantages over the other in such areas as leadership, training, doctrine, and weapons. Precisely for this reason, many famous battlefield commanders throughout history (e.g., Napoleon and von Manstein) have not hesitated to attack in the face of equal or even greater numbers. Offensive (or defensive) maneuver did not always provide them victory. But often it did.

If history is an indicator, the risks posed by an offensive campaign of maneuver are important to a modern NATO defender contemplating a linear employment doctrine, even in a situation of parity. A linear defense is not automatically doomed to failure against an offensive maneuver strategy, but neither is success ensured. Rather, the outcome of a clash between the two strategies at parity is problematic. A variable can be affected by the interplay among several determining factors, including the defender's size in relation to the terrain being protected. In gauging the feasibility of linear defense at low NATO force levels, the central issue is: Can the attacker concentrate sufficient forces at locations along the defender's front line to foster and then exploit a major breakthrough, permitting decisive maneuver in the defender's rear? Or can the defender use the local forces and classical advantages at his disposal to control, halt, and ultimately rebuff the attacker's effort?

This issue emerges from a complex sequence of offensive-defensive interactions that shape the battlefield at low force levels and set the stage for the clash between an offensive maneuver strategy and a linear defense. This analysis will first define in greater detail exactly what is meant by a "linear" defense. Then it will examine offense-defense interactions at a force level of 30 divisions apiece, a fairly low concentration of forces in present-day terms for 750–800 kilometers of frontage. The Alliance's CFE I position, for example, calls for NATO to retain enough equipment there to support about 40 divisions.

The Theory of Linear Defense

In its purest form, "linear defense" can be defined as a theater employment doctrine that leads a defender to place highest priority on building a uniformly strong front wall along his entire exposed border. As a result of this preferential pattern, only a small percentage of the defender's total forces (e.g., 25 percent or less) are withheld as "theater reserves," to the rear of the forward corps sectors. Moreover, the defender plans to maintain this front wall

generally intact throughout the initial stage of operational maneuver, or long enough for the attacker to play out the first phase of his offensive strategy. The defender's doctrine thus is linear in space and positional in time. It leads the defender to allocate a fixed segment of his force posture to populate the front wall in reasonable density. By "reasonable" is meant sufficient density to establish a contiguous front line that cannot readily be breached, with weapons sufficiently close to each other to provide a fully integrated combined arms defense. Only after this thick front wall is established are any remaining forces formed in the rear areas, where they can be employed as operational reserves.

In this employment doctrine, the number of forces assigned to the front line is a constant, while the number withheld as reserves is a variable that would decline as the defender's posture shrinks. Table 18 displays how a linear defense with 30 divisions might be constructed along a 750 kilometer front in Central Europe today. The defender's posture is divided into two Army Groups (A and B), each composed of four corps of three divisions apiece. Some 24 divisions are committed to the front wall, and only six divisions are withheld as corps or Army Group reserves. Although the frontages of the frontally committed divisions might differ somewhat from one to the other, the theater wide average is about 30 kilometers. This frontage is similar to what has prevailed in NATO's employment doctrine in recent years.[3]

This employment doctrine would leave the defender with 80 percent of his forces committed on the front line; the percentage would rise if the defender's posture was larger, but it would decline to zero if the posture shrank to 24 divisions. The defender, of course, has other options available. He could dispense with a front wall entirely by concentrating his forces, enabling him to maneuver in mass against the attacker and possibly conduct offensive operations of his own.

Several factors might lead him to elect a linear defense provided he has sufficient forces (24 divisions) to populate a front wall. This employment doctrine is easier to execute than a maneuver defense, and it helps ensure that the attacker will not easily seize exposed territory along the frontier. Also, a linear defense enables him to employ the classical advantages of being on the defense, including preselecting the terrain for engagements, constructing prepared positions, and fighting a prepared set-piece battle that enables him to make optimal use of all the elements of his combined arms team. The greater the time to make these preparations, the more able is the defender to enlist all these advantages on behalf of his cause.

Table 18

Employment Pattern of a Linear Defense with 30 Divisions

	Army Group A				Army Group B			
Corps sectors	A	B	C	D	E	F	G	H
Divisions	3	3	3	3	3	3	3	3
Theater reserves			3				3	

[3]See Department of Defense, 1984.

The analysis assumes that the defender's employment doctrine does not completely freeze him into a positional defense. Although the defender would begin a battle with his forces deployed according to Table 18, he would be free to maneuver his forces within each Army Group to different locations. Also, he could shift the reserves of one Army Group to the other, but he would not plan to shift the frontally committed forces of one Army Group to the next. Thus, the most force that he could assemble in any single Army Group is 18 divisions. As defined here, this restriction on large-scale operational redeployment is a distinguishing feature of a linear defense. It is admittedly a somewhat artificial rule that might well be abandoned in actual practice, and that makes linear defense appear to be an easily knocked-down strawman. But this rule, artificiality aside, allows the analysis to bring linear defense into sharp relief and to examine its theoretical properties in a fixed, specifically defined setting.

Although this analytical rule seemingly imposes a stiff constraint on the defender's employment options, a variety of practical factors enhance its realism. Historically, defenders often have not chosen to undertake large-scale redeployments when they have sensed a concentrated attack coming their way, or even during the first critical days of battle. The theory of linear defense, as defined here, does not base its analysis of force needs on the assumption that units in less threatened areas of the battlefield are highly fungible.

To construct a strong frontal wall, the bulk of the defender's forces preferably should be committed along the entire front line well in advance of the war's breaking out. A well-developed network of forward positions with ample depth takes several days to prepare. Since it should be fully completed before battle begins, defending units must begin preparing it before enemy forces are committed to battle. These units cannot complete this task if they are still moving toward their positions on D-Day.

The attacker's employment strategy is often not sufficiently discernible to allow the defense much flexibility. As attacking forces approach the battlefield, their physical disposition can signal to the defense where a concentrated blow is likely to fall. In theory, this signal can allow the defense to begin counter-concentrating in advance. In reality, however, the situation normally would be more complex, ambiguous, and confusing. A skillful attacker could first feint in one direction and then another before committing himself, preventing the defender from knowing where the main blow was about to be launched. Beyond creating uncertainty, these feints might be designed to manipulate the defender into making mistakes, for example, by moving those forces that should be digging in while positionally committing forces that should be moving. By assigning first priority to constructing a solid front wall regardless of enemy lures, the defender can avoid this risk while not being diverted from laying the foundation of his defense plan. However, the defender thereby unavoidably disqualifies a large portion of his forces from the early redeployment equation.

As the battle begins unfolding, the defender's options broaden. Even then, however, other factors can constrain the defender's ability to redeploy quickly, decisively, and effectively. One factor is a potential failure in intelligence. The defender might not correctly interpret the attacker's intentions and movement patterns. This constraint was particularly important in earlier years, but it is becoming less prevalent now as modern intelligence systems become increasingly sophisticated. Another reason is timing and relative motion dynamics. Especially if the attacker enjoys interior lines of communication that give him less distance to cover, he might be able to converge at any single point faster than the defender. Furthermore, because of confusion and conflicting internal priorities, the defender might be unable to quickly reach a decision to move laterally.

Cross-theater movement can be a difficult undertaking. Forces transiting the rear areas can be delayed by air strikes and by enemy airborne and air assault operations. Particularly troublesome is the task of withdrawing forces that are committed on the frontline, even from sectors where enemy troops are not heavily concentrated. During withdrawal, these defending forces would become exposed to enemy indirect fire and might suffer heavy casualties. Also, the process of withdrawal can create temporary gaps through which nearby enemy forces could pass, potentially in enough strength to disrupt the defender's rear areas. All this can delay the redeployment process considerably.

An often overlooked constraint is logistics. Combat forces can be moved fairly quickly, but more is involved than transporting them alone. Support forces also must be moved, particularly when the defender's support structures from one area lack the capability to accommodate reinforcements from another area, or when the defender's posture is a multinational coalition that lacks logistic interoperability. In this case, redeploying combat forces would have to bring their own logistic support units, and these units are normally large and cumbersome. Even if the road and rail network is adequate, the process of deploying them often can be time-consuming.

For all these reasons, a goal of a linear defense, as defined here, is to create a state of affairs whereby the defender is not heavily reliant on theaterwide redeployment tactics at the onset of hostilities. By its nature, this doctrine permits the attacker the advantage of concentrating, and it is skeptical about the defender's ability to match him. The attacker knows his precise military plans in advance and can orchestrate events accordingly, whereas the defender is faced with major uncertainty. If this uncertainty leads to poor force-allocation decisions, the defender can wind up with his forces becoming disorganized, widely separated, and unable to fight together coherently against a cohesive enemy, the very vulnerability that an offensive maneuver strategy seeks to foster. A defensive strategy of early counter-maneuver does not make that chaotic situation inevitable, but it increases the risk that it will come to pass.

Linear defense consciously trades off forward strength at the expense of early counter-maneuver. Ideally, it aims to render the defending forces in each Army Group strong enough that, on their own, they can absorb and contain a concentrated enemy assault. It applies this goal to the entire theater, not just to any one sector that seems most likely to be attacked. And it aspires to build this capability before the onset of hostilities, rather than attempting to create it after war has begun.

This theory does not reject the idea of both bringing reserves forward and transferring frontline forces laterally to the point of main contact after an invasion has started. But this reliance is limited to forces that are located nearby or at least are quickly available. The option of a larger theaterwide redeployment is put off until after the enemy's initial assault has dissipated and the defender begins preparing to launch a major counterattack. In terms of the general model used here, this means that Army Group A's forces should be sufficiently strong not to be dependent on the timely arrival of those Army Group B forces that are originally deployed in forward positions, a long distance away.

This concept, it must be emphasized, is a planning goal. It postulates a design standard and a test of sufficiency for gauging how a pure linear defense ideally should be prepared to reduce risks to a low level. Decidedly, it is not a rigid prescription for actual defensive operations. If events were to take an undesirable course in a war, clearly the defender would try to improvise. But for the moment, the analysis will proceed on the assumption that the defender initially would maintain this employment doctrine even in the face of adversity.

The purpose here is analytical. By holding this employment doctrine constant, it is possible to gauge the risks of maintaining a linear array as the defender's forces are reduced.

Linear Defense Against a Broad Front Attack

Three broad attack options will be examined ranked according to the degree of stress that they would place on the defender's linear employment doctrine. All attacks are assumed to be mounted at a condition of parity.

- A broad front offensive strategy that places the least stress on a linear defense.
- A concentrated attack strategy aimed at seizing portions of the defender's territory. Compared with the other two strategies, this strategy places a medium level of stress on the defender.
- A maneuver strategy aimed at inflicting operational defeat on the defender's forces by fostering a breakthrough and then exploiting it in the defender's rear areas. This strategy places the greatest stress on a linear defense.

Because a maneuver strategy creates the highest risks for the defender, it offers the attacker the highest payoffs of these three strategies. The analysis will concentrate on this strategy and the problems it raises for a linear defense at low force levels. First, however, a brief review of the other two strategies is in order.

A maneuver strategy is not necessarily the attacking army's preferred choice. It is a complicated, demanding, and risky strategy that can leave the attacker exposed to defeat if it is not executed effectively. By contrast, the least demanding strategy of a broad front attack has attractive advantages provided the attacker is strong enough to overpower the defender without concentrating. Table 19 displays how an army composed of 30 combat divisions might conduct a broad front attack. In this scheme, 24 divisions are deployed on the attacker's front line along a linear array. The remaining six divisions are held in reserve; they are available for commitment, but only to strengthen the front line, not as a concentrated force for breakthrough and deep maneuver. The attacker's employment doctrine thus is a mirror image of the defender's scheme. As a result, not only is the theater force ratio one of parity, but the engaged ratios within each corps sector are also 1:1.

A broad front campaign plan simply requires participating attacking formations to advance in linear fashion abreast of each other. Complicated maneuvers requiring synchronized coordination therefore are unnecessary. Also, a broad front attack is supported logistically. The road network supporting the advancing army is less likely to be overloaded in

Table 19

Broad Front Attack Strategy vs. Linear Defense

	Defender Army Group A				Defender Army Group B			
Sectors	A	B	C	D	E	F	G	H
Attacker reserves			3				3	
Attacker frontline divisions	3	3	3	3	3	3	3	3
Defender frontline divisions	3	3	3	3	3	3	3	3
Defender reserves			3				3	

any one sector because combat and support units are distributed uniformly across the battlefield. Rear area support units along the advancing front, in turn, are required to render a constant level of support to a limited number of local combat units. They are not compelled to intensify and possibly overload their efforts as additional forces are concentrated in their sectors. Equally important, advancing combat units never expose their flanks to enemy counterattack. A broad front attack therefore entails fewer operational risks than a maneuver strategy that uncovers its own flanks in an effort to attack the defender's flanks. Finally, it leaves no exposed gaps in its own line, and it can quickly swing over to the defense if the enemy launches its own attack.

A broad front campaign offers the attacker an array of political-military options. The attacker can choose to advance only a limited distance into enemy territory and then halt and dig in, thereby inflicting a political defeat on the defender. The attacker can continue his march, conquering territory, cities, and other targets of value as he advances. One drawback is that it generally unfolds in a slow and ponderous way. Although it places a constant level of stress on the defender's posture, it exerts no concentrated pressure against any individual weak links. As a result, it normally does not offer a quick route to decisive victory. But if the attacker is strong enough to continue the advance, it offers a safe and feasible way to victory by steadily wearing down the defender and driving him backward. Precisely for these reasons, General Eisenhower mostly pursued a broad front campaign as the allied armies advanced across the West European theater in World War II. Eisenhower was subjected to a barrage of criticism from British Field Marshal Montgomery and even U.S. General George Patton, both of whom preferred a more adventurous strategy of concentration, breakthrough, and maneuver. But his decision ultimately proved a wise one. Within a year of landing at Normandy in June 1944, his allied forces defeated the German Army and advanced deep into Germany itself, thereby helping bring down Hitler's regime.

From the attacker's vantage point, the drawback of a broad front strategy is that it normally requires a sizable numerical superiority to execute in a sustained way. This strategy worked for Eisenhower because allied forces enjoyed a sizable preponderance over the German forces for most of the campaign and had major advantages in replenishment and logistic sustainability. A battle fought at parity, however, creates a decidedly different set of circumstances that seemingly would work in favor of a linear defense.

Were this war to be fought when the defender was unable to make any advance preparations, the outcome, assuming absolute equality on both sides, would probably be a stalemate. The attacker initially might advance slowly into the defender's territory. A rate of about five kilometers per day seems a reasonable estimate since 1–2 local enemy attacks could be launched per day across the battlefield and purely tactical considerations would lead the defender at times to withdraw a short distance. But both sides would suffer attrition at about the same rate. Eventually, the attacker's advance would lose energy, and the two armies, both battered but still equally large, would find themselves facing each other a short distance into the defender's territory.

This picture would change dramatically, however, if the defender had the time and opportunity to make preparations. The main effect of preselecting the terrain, preparing positions, and taking related measures is to enhance the firepower lethality of the defender's weapons while reducing their vulnerability. This important change would transform the attrition battle from an equal contest into a one-sided war favoring the defender. Table 20 makes plausible assumptions about the effect that advance preparations have on the attacker's and defender's lethality. Their net effect is to transform the attrition struggle in favor of the defense. These assumptions, which postulate a 50 percent change in lethality for

Table 20

Effect of Defensive Preparations[a]

	Without Advance Preparations	With Advance Preparations[b]
Attacker lethality	.05	.075
Defender lethality	.05	.025
Lethality ratio	1:1	1:3

[a]"Lethality" is defined as the daily capability of a ground division to inflict losses (measured in percentage of total DEFs) on a single enemy division, assuming an initial engaged force ratio of 1:1.

[b]"Preparations" are defined as prepared positions through the defender's divisional rear areas, supplemented by fire-covered obstacles that canalize the enemy advance.

both sides, are consistent with data commonly employed in RAND and DoD dynamic computer models. They are derived judgmentally; if anything, they underestimate the advantage accruing to the defender, especially if the attacker fails to suppress the defender.

Although the effects on the battlefield would not be particularly noticeable after the first day of combat, they would become increasingly apparent in succeeding days. Table 21 displays the attrition effects after five days of combat. In the case where the defender makes no advance preparations, both sides are reduced to 77 percent of their original combat power, the force ratio is still 1:1, and the attacker has advanced 50 kilometers. In the case of advanced preparations, the attacker has been reduced to 64 percent of his original strength, while the defender still has 90 percent of original capacity. The force ratio thus has changed to 1.40:1 in the defender's favor. Also, the attacker, slowed by adverse attrition and a mounting unfavorable force ratio, has seen his advance reduced to about 30 kilometers. Stated more simply, the attacker finds himself losing the war because of unfavorable attrition

Table 21

Effect of Defender Preparations on Combat Outcomes on D+5

	Without Advance Preparations	With Advance Preparations
Attacker strength remaining (percent)	77	64
Defender strength remaining (percent)	77	90
Force ratio	1:1	1.40:1[a]
Attacker advance (km)	50	30

[a]Force ratio favors the defense.

dynamics brought about by the defender's preparations. Bogged down only a short distance into enemy territory, the attacker has lost one-third of his forces and now faces a larger defender who has lost little of his original strength.

Were the attacker to continue his assault, these combat dynamics would accelerate even more decisively in the defender's favor. The defender's advantageous force ratio and lethality advantage would translate into an increasingly predominant force ratio. Armored battles of this nature commonly are modeled with some version of the Lanchester Square equation, as displayed in Table 22.[4]

Figure 38 shows the results of this attrition battle fought to the bitter end as modeled by Lanchester Square. It suggests that the battle would end with the attacker completely destroyed and the defender still possessing about 80 percent of his original strength. In particular, once the force ratio exceeds about 3:1, an explosive attrition dynamic ensues that results in the rapid destruction of the attacker with little additional damage to the defender

Table 22

Lanchester Square

$$L(MD_0^2 - MD_n^2) = L(MA_0^2 - MA_n^2)$$

L equals the lethality of each side's units.
MD_0 is the defender's size at the start; MD_n is its size at the finish.
MA_0 is the attacker's size at the start; MA_n is its size at the finish.

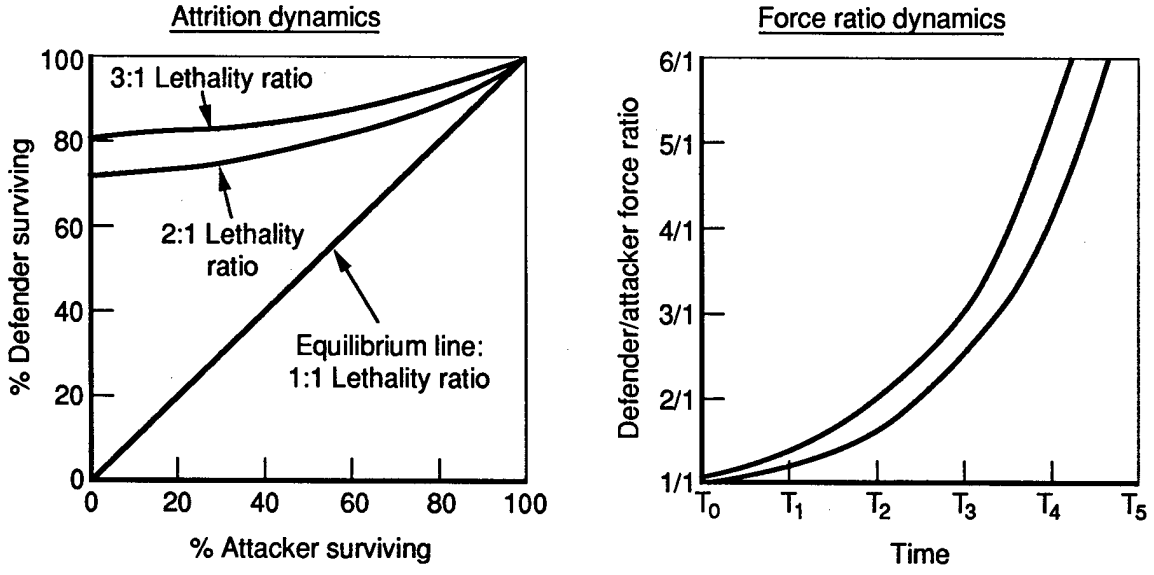

Fig. 38—Combat Dynamics Over Time with Lanchester Square

[4]Taylor, 1981, pp. 20–83.

(or additional loss of territory). In essence, the defender's larger size as the battle wears on, coupled with his unchanging 3:1 lethality advantage, eventually produce a complete victory, provided the attacker does nothing to halt the battle or reverse his fortunes. Long before the attrition battle has ended, the defender would enjoy a sufficient force advantage to swing over into a counterattack, regain lost ground, and destroy the aggressor's remaining forces in the process. This outcome stands in sharp contrast to the results of an attrition battle fought without defender preparations, which yields a constant force ratio and simultaneous mutual exhaustion.

These results, it must be emphasized, are illustrative. The Lanchester Square equation is derived from a set of differential equations anchored on a single conceptual model postulating that the interaction of numbers and lethality is the chief determinant of military outcomes. The equation shows how one side can optimally employ a fairly small initial edge in either quantity or quality to achieve a decisive victory. In a sense, it helps illustrate how the initiative can be seized and physically exploited. Its drawback lies in its singular focus on optimal performance and its failure to allow for variations in tactics and doctrine that could alter the underlying model. It thus portrays the initiative and effectiveness as being one-dimensional, deriving solely from mass, weapons, and technology.

Lanchester Square mathematics originally were developed to analyze 19th century naval engagements in which line formations of battleships attacked each other in a stylized manner, reflecting maritime doctrine of the time and resembling this conceptual model. The equation itself grew to prominence in the 1920s when F. W. Lanchester applied it to air-to-air engagements among a limited number of combatants performing along this model's lines. For these engagements, small differences in numbers or quality, in fact, did promise to produce a decisive effect on military outcomes. Because of the cascading effects of an initially small firepower edge, for example, five battleships were capable of completely destroying four equal opponents. Likewise, acquisition of larger guns (e.g., 14 in. rather than 12 in. tubes) threatened to transform a numerical balance into a decisive advantage for the side engaged in modernization. For these reasons, and because doctrinal alternatives were limited, Lanchester Square was an appropriate model for studying naval and air engagements.[5]

In pure form, Lanchester Square's ability to model modern, sustained ground warfare between large armies, employing many different weapons and tactics over a long time period, is questionable. Ground combat is more complicated and less rigidly deterministic than Lanchester Square's underlying conceptual model. In particular, as Stephen Biddle and Joshua Epstein have pointed out, Lanchester Square is a mathematically unstable equation in which minor perturbations increasingly produce major changes, and do so in a mechanical way that seems divorced from the ebb and flow of real combat. These and other critics contend that it exaggerates the effects of marginal differences in both size and lethality. They conclude that it underestimates the capacity of military forces to make intelligent adaptations that can help restore equilibrium to combat.[6]

Several alternative equations, endeavoring to reflect real-world battlefield operations, have been developed that suppress these unstable effects by diluting the advantage of larger numbers and greater lethality. The Lanchester Linear equation (normally used to model indirect fire) is one example, and there are others. Applied to this case, these equations would show that the defender would win less decisively than Lanchester Square implies. But

[5]See Hughes, 1986.

[6]See Epstein, 1987. Also, see Biddle, 1988. Epstein's adaptive model allows the defender to make intelligent adjustments to an adverse situation, thereby restoring equilibrium to the battle. The RSAS system is designed to allow both sides to make adaptations through use of terrain, tactics, and other factors.

as long as they grant the defender substantial qualitative advantages from prepared positions and preselected terrain, nearly all would still show a similar final result. Lanchester Linear, for example, forecasts that 66 percent of the defender's forces would survive, rather than 80 percent under Lanchester Square, with the attacker reduced to zero in both cases.

Specific numerical forecasts aside, a well-prepared army conducting a linear defense can decisively defeat an equally large attacker advancing in an unimaginative, broad front way. Indeed, unless the classical advantages of fighting on the defensive are stripped away, a well-prepared defender can realistically fight outnumbered and still hope to contain a larger attacker. The exact threshold of manageability is a function of analytical assumptions about lethality and force interactions as portrayed in the mathematical equations that are used. The important variable is the defender's ability to extract sufficiently favorable exchange ratios to maintain at least a constant force ratio, thereby preventing the cascading dynamics of attrition from starting to operate against it. The Lanchester Square equation postulates that to achieve an attrition stalemate (both sides reach zero simultaneously), an outnumbered defender must achieve a lethality advantage that is the square of the original force ratio. Conversely, this equation also postulates that an outnumbered defender holding a given lethality advantage can manage a numerical disparity that is the square root of this advantage.

With this equation and the lethality data here, a 30-division defender would be capable of achieving an attrition stalemate against an attacker about 1.73 times as large, albeit at the cost of steadily losing ground until the attack had lost its steam. Lanchester Linear, which more heavily favors the outnumbered side by postulating a manageable disparity equal to the defender's lethality advantage, forecasts a stalemate against an attacker 3.0 times larger. Less favorable assumptions would result in a lower threshold of defensive tolerance; the converse is also true. Regardless, the central point stands out: A defending army large enough to construct a solid, well-prepared forward wall along its borders is not easily dislodged by a crude frontal assault by anything less than a much larger force.

History broadly supports this judgment. The stalemate phases of World War I and the Korean War are classic cases in point. Another good example is the German defense of the west wall in World War II. For nearly three months in late 1944, the German Army, despite being heavily outnumbered, held the allied armies at bay. In particular, the allies suffered bloody losses and gained little in their efforts to advance through the dense Huertgen Forest. The allies finally were able to resume their advance in January 1945, but only after the Germans had squandered much of their strength in their ill-fated Ardennes counterattack in the Battle of the Bulge. Had the Ardennes offensive not been launched, the allied advance probably would have been delayed longer still.[7]

Military circumstances have changed considerably since World War II. The net effect of modern technology, however, appears to have enhanced the defender's advantage in this kind of attrition slugfest. For example, 30 years ago, high-speed armored assaults, employing tanks and armored personnel carriers, were effective tactics for quickly closing on, isolating, and destroying defensive positions. Modern tanks can now fire effectively to a range of 2000 meters and beyond; earlier models could reach only to about 1000 meters. Consequently, enjoying stationary positions behind prepared defenses that leave only their turrets vulnerable, they can destroy several exposed attacking tanks in a single engagement. Lethal, long range antitank missiles mounted on infantry fighting vehicles have further helped the defense. These weapons are less helpful to attackers because moving forward rapidly over

[7]See Eisenhower, 1986, pp. 482–553.

bumpy terrain reduces their accuracy. As a result of both developments, defending units now are better able quickly to destroy attacking enemy armor before it closes on them.[8]

The effects of modern weaponry have also been felt in areas other than armor. Modern artillery, with advanced area munitions, logistic support, and fire control systems, has become quite lethal against attacking formations caught in the open. Defending units, under cover, are less vulnerable. Scatterable mines have enabled the defender to sow minefields quickly; modern combat engineers can promptly create other obstacles that can be covered by direct and indirect fire. Finally, tactical air forces have acquired the avionics and munitions to conduct effective close air support strikes against exposed targets. These developments have helped the attacker to some degree, but their overall effect has been to swing the technology balance in the defender's favor.

In summary, since early this century the strategy of a broad front attack against a prepared linear defense has made little sense in a situation of parity, or anything like it. In Central Europe, the Soviets might have contemplated this strategy in earlier years when they enjoyed a large preponderance (2:1 or more) and NATO lacked enough forces to build a strong wall. But NATO now possesses sufficient forces for this purpose and the Soviets are facing a balance of approximate parity in the years ahead. For them, as well as for attacking armies in general, parity in the modern era forecloses the option of frontally assaulting a well-prepared linear defense.

Linear Defense vs. a Concentrated Attack

The daunting prospect of assaulting the main strength of a prepared linear defense easily could lead the attacker to seek a potentially more lucrative approach. This stage of the offense-defense interaction involves a decision by the attacker, as his forces approach the battlefield, to quickly concentrate his units at selected points along the defender's front line, while thinning out elsewhere. Table 23 illustrates the effects of a strategy in which the attacker concentrates 24 divisions on the left side of the battlefield to achieve a 1.60:1 advantage against the defender's forces of Army Group A there. On the right side of the battlefield against Army Group B, meanwhile, the attacker is left facing a 2.0:1 disadvantage along the line (2.5:1 if the defender's immediately available reserves are counted). Presumably he hopes that the defender will not swing over to the offensive on the right while he himself is attacking on the left. Alternatively, the attacker might hope that his five divisions can temporarily contain a defender advance on the right or possibly draw the defender further out of position for responding to developments on the left.

Table 23

Theaterwide Effect of Attacker Concentration Tactics (divisions)

	Army Group A	Army Group B
Attacker	24	6
Defender	15	15
Ratio (A/D)	1.6:1	1:2.5

[8]See Deitchman, 1983.

Table 24 illustrates how the attacker's forces opposite Army Group A might be employed in executing a concentrated attack strategy aimed at maximizing the attacker's conquest of enemy territory. To achieve this goal, the attacker concentrates 18 of his 24 divisions directly opposite two of the defender's four corps sectors, manned by three divisions apiece. Terrain and other constraints would prohibit the attacker from narrowing his own division frontages lower than 12–15 kilometers apiece. As a result, the attacker's forces advance in two waves of echeloned formations. This employment scheme gives the attacker an initial 3:1 edge in the two concentrated sectors but only a 2:1 edge in actually engaged forces.

In this strategy, the attacker's operational plan would endeavor to drive forward in Sectors B and C. The attacker's forces in other sectors also would advance to the extent possible, primarily hoping to inhibit the defender's units from redeploying to the more threatened area, but they would not attack as aggressively. In Sectors B and C, the attacker might hope to wear down the defender's forces in an attrition fight, but attrition would be merely one means to a larger strategic end. His main goal would not be the operational destruction of the defender's entire posture but to seize territory, possibly to inflict a political defeat on the defender, while avoiding any risk of the destruction of his own forces. At the end of this campaign, the defender might retain combat formations still capable of fighting, but he would have been pushed well back into his own territory and presumably would lack the strength to launch a counterattack. Having accomplished this goal and demoralized the enemy politically, the attacker then might resort to diplomatic means to end the war.

The attacker would pursue this goal presumably by advancing only straight forward. The forces in Sectors B and C would advance as rapidly as possible, hoping that offensive pressure there would compel the defender, fearing to expose a flank, to pull back along the Army Group A's entire front. But even if an exposed flank came open, the attacker would not pivot and move laterally or diagonally in an effort to envelop the defender's forces in nearby sectors. By driving only forward, but as deeply as possible, the attacker might aspire to translate his 3:1 edge into an initial advance rate of 10–15 kilometers per day, or more. His ideal goal might be to advance 200–300 kilometers before running out of steam; minimally, he might aim for a 100 kilometer advance.

The defender's goal would be to contain this attack in the forward areas, ideally within 30 kilometers of the border. His initially available forces would be too small to achieve this

Table 24

Force Deployments vs. the Defender's Army Group A

Sector	A	B	C	D
Attacker				
Reserves		3	3	
Front line	3	6	6	3
Defender				
Front line	3	3	3	3
Reserves		3		

objective; left on their own, they might be overpowered since they are outnumbered. Within a few days, however, Army Group A's reserves (three divisions) and two or three divisions drawn from adjacent sectors would begin converging on the scene. Somewhat later, Army Group B's reserves (three divisions) might also arrive. These forces would reduce the local force ratio (attrition aside) to 1.50–1.30:1. While hardly parity, this is a theoretically manageable balance as long as the defender continues to benefit from prepared positions.

To what degree would the attacker succeed in attaining his goals, taking into account Army Group A's use of these reserves? This question will be discussed below when this section examines alternative NATO force levels with dynamic analyses that can assess the interactions. Suffice it to say that the answer depends heavily on whether these reserves arrive in a timely fashion. "Timely" is measured by whether these reserves arrive fast enough to help maintain an adequate mass of forces in front of the attacker's advancing columns, taking into account attrition dynamics on both sides. The defender would constantly need enough forces to maintain a competitive stance in the attrition process. Also, he would need to slow the attacker's advance down so that his own combat forces do not retreat faster than combat engineers can prepare successive layers of positions in the rear areas. Without a long period of strategic warning, these positions normally are prepared in advance only to about 25 kilometers, but the depth can be extended after fighting begins if combat engineers are given sufficient time.

With a timely arrival, Army Group A would probably be able to inflict adverse exchange rates on the attacker, halt his advance, and compel him to swing over onto the defensive. With the attacker bogged down only a short distance across the border, the defender, in the extreme case, might then be able to mount a counterattack aimed at restoring the original front line. Without a timely arrival, the defender, outnumbered and lacking prepared positions to compensate for this disadvantage, would face less favorable prospects. Indeed, the attacker might be able to swing the attrition battle in his favor, maintain a rapid advance, and drive deep into the defender's territory. Timeliness, along with the absolute number of defender reserves, thus is a critically important variable.

Linear Defense vs. a Maneuver Strategy

Both the broad front and concentrated strategies would reflect limited ambitions by the attacker and would place correspondingly modest stress on the defender's linear array. A maneuver strategy that reflected higher attacker ambitions would, if executed effectively, place greater stress on the defense. The attacker's goal in this strategy would be to inflict an operational defeat on the defender. Seizing terrain would be a secondary issue, a byproduct of a successful outcome rather than an immediate goal. Essentially the attacker would aspire to inflict a high degree of physical destruction on the defender's forces, rendering them incapable of continuing the war and compelling them to surrender.[9]

The attacker would not attempt this task through the mechanistic application of superior mass and firepower growing out of a broad front attack against the defender's linear defense. That course of action would be precluded by the defender's equal strength at parity coupled with his prepared positions in the forward areas. Instead, the attacker would employ aggressive maneuver operations in an effort first to gain breakthroughs that would

[9]For analyses of manuever strategy, see Liddell Hart, 1967; Simpkin, 1985; von Mellenthin and Stolfi (with E. Sobik), 1984; von Manstein, 1984.

strip away the defender's positions and dislocate his maneuver units.[10] Success in this endeavor presumably would enable the attacker then to engage the defender's units in detail and destroy them in a series of violent meeting engagements in the rear areas.

Although various employment strategies are possible, this analysis will concentrate on a "single envelopment" of the defender's forward units. It is the offensive plan that attacking armies seem most often to have pursued in a situation of parity. It calls for the attacker to gain a major breakthrough in one sector of the defender's front line and then to conduct a flanking maneuver in the rear. An alternative concept would be a two-pronged "double envelopment." Although it could be pursued in a situation of parity, it would place greater demands on the attacker's forces and consequently is most commonly associated with battles in which the attacker has a large advantage. For simplicity's sake, this analysis will examine the single envelopment strategy in most detail. The conclusions drawn, however, are broadly generalizable to both. These two campaign plans are schematically displayed on Fig. 39.

Table 25 displays how the attacker's forces might be committed in pursuit of this single envelopment breakthrough strategy. Fully 18 of the attacker's 24 divisions are concentrated directly opposite Sector D on the defender's line. These forces are arrayed in three tactical echelons of six divisions apiece, giving the attacker a total force ratio advantage of 6:1 in this sector, although only a 2:1 edge in engaged forces. Notice also that in this scheme, the attacker has executed economy of force tactics by deploying only two divisions apiece against each of Army Group A's remaining three sectors. It is this tactic, coupled with the focus on only one main sector, that enables the attacker to drive up his local force from the 3:1 edge of the concentrated strategy to a 6:1 advantage now.

The attacker's initial operational goal would be to gain major breakthroughs in this main sector by driving through defending units there. Attrition would be part of his battle

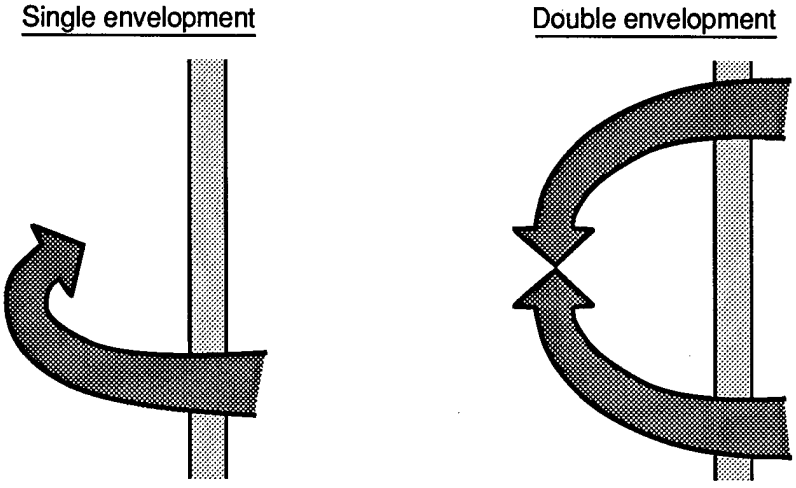

Fig. 39—Maneuver Strategy: Single vs. Double Envelopment

[10]A "breakthrough," as defined here, is a sufficiently large rupture of the defender's front line to allow the attacker to advance into the defender's rear areas in multidivisional strength. The term does not refer to minor penetrations that can be readily closed through tactical adjustments by the defender.

Table 25

Force Commitment Profile for a Single Envelopment Breakthrough Attack

Forces	Sector A	B	C	D
Attacker				
Echeloned reserves				12
Front line	2	2	2	6
Defender				
Front line	3	3	3	3
Reserves		3		

plan, but his immediate concern would not be to employ raw firepower to destroy the defender's units along the way. Instead, the attacker's forces in the first assault echelon would aim to disrupt and bypass these units and then advance rapidly through the depth of the defender's positions. The attacker would aim to create opportunities for this deep advance by pushing aside individual defending battalions and brigades at critical locations—for example, at the seams of divisional and corps boundaries and along key roads. These tactics would create open avenues, gaps, and corridors for an unhindered advance by large armored reserves coming from the second echelon. The attacker thus would endeavor to pry open the defense rather than bludgeon it to death, relying heavily on synchronized combined-arms operations in which artillery and mechanized infantry, supported by helicopter and airborne forces (along with tactical air power), would pave the way for high-speed tank spearheads aimed at punching through.

These operations would be expedited by agility, skillful tactical maneuver (e.g, a coordinated series of diagonal advances), and exploitation of the terrain. Also playing an important role would be an effort to accelerate the momentum of battle to the point where the defender is left unable to respond rapidly enough. The attacker would hope to inflict enough confusion and uncertainty on the enemy's commanders to paralyze them into indecision, while using shock action to break the morale and psychological cohesion of their troops. The attacker's breakthrough campaign would be anchored on the calculation that combat power is more than a function of mass and maneuver, that expert planning, teamwork, velocity, tempo, and sheer willpower are important, potentially synergistic, force multipliers.

The attacker's units in nearby secondary sectors would simultaneously endeavor to employ heavy indirect fire and aggressive probing actions to pin down enemy troops opposite them. The goal would be to prevent these troops from falling back. Success in this effort would foster a situation in which large attacking forces are driving deep into the rear areas, while the bulk of the defender's forces are still tied down well forward. Having gained exploitable military leverage over Army Group A's entire posture this way, the attacker would have set the stage for the next act in his unfolding play.

Within the main sectors, the attacker's armored formations, driving rapidly along multiple axes and thrust lines, would advance into the enemy's rear areas. They would hope to have achieved this breakthrough faster than Army Group A's reserves can have moved to both main sectors and begun mounting organized opposition. In that event, the attacker's

forces would execute turning movements and begin fanning out left and right. A few attacking units would probably establish a defensive screen to block Army Group B units from quickly redeploying to the scene. With the battlefield isolated, some attacking forces would pivot at a fairly shallow depth and launch strikes against the artillery, command posts, communication center, and support structures of nearby defending combat units. Other attacking forces would pivot deeper and execute meeting engagements to defeat any defender reserves in their way, completing their single envelopment.

If successful, this breakthrough and envelopment operation would place nearly all of Army Group A in a disadvantageous position. Army Group A, originally bonded into a cohesive whole, would have been fractured and dislocated. Its combat divisions would have been divided from one another and either atomized or broken into small clusters incapable of readily supporting each other. Many of these divisions would have been driven off of their prepared positions; those that remained would find themselves exposed to attacks from vulnerable flanks and the rear. Many also would have been cut off from their artillery, ammo stocks, POL depots, and service support forces. Further, most of Army Group A would have had its arteries of communication and lines of retreat severed. Army Group A's surviving forces would still be capable of fighting, provided their morale remained intact, but their combat power and sustainability in many cases would have been appreciably reduced. Equally important, Army Group A would now have difficulty fighting in a coordinated way. Separated, broken apart, and weakened, its forces would now be vulnerable to being defeated one by one.

This portrayal, of course, merely describes a concept of operations. It does not demonstrate that the attacker, at these force levels in this case of parity, could successfully carry this operation out. Many imponderables exist, and the attacker's prospects would hinge on many variables. One critical issue would be whether the attacker, with 1.6:1 local edge against Army Group A, would have enough of an advantage to carry out all the demanding phases required in this plan. Another would be whether the attacker's forces could muster the military skill to execute what is clearly a demanding and risky operation. Would Army Group A's frontline forces at the main points of attack buckle and be swept aside as totally as the attacker would wish? Could defender bring enough reserves to bear, locally and from Army Group B, to stem the tide? The question of reserve availability—in size and timing—thus arises again, in an even more important way.

In the absence of a more detailed dynamic analysis, these questions cannot yet be addressed in an insightful way. Clearly, a range of outcomes is possible, including defensive success. Especially if the defender were to employ his frontline forces more effectively or marshall larger reserves in time, the attacker might well be rebuffed, or at least contained. At the same time, three considerations—all measurable—suggest that a decisive attacker victory would be far from implausible. The attacker would have enough forces available to conduct a serious offensive campaign of breakthrough and maneuver. A force of 20–24 divisions is the size of a standard Soviet "front," which is designed for operations of this scope. Furthermore, some time would be required before sizable reserves from Army Group B could become available. A single division, if unhindered, might arrive within two or three days, but a full corps, with support forces, could take a week. Larger forces, if available, would take even longer, giving the attacker a time window within which to operate. In this period, the attacker might be successful in transforming the battle into a widespread meeting engagement in which the defender had no access to prepared positions. The defender would have to marshall more than two or three extra reserve divisions from Army Group B. In principle, a

1:1 ratio would be needed. Depending on attrition, up to nine additional divisions could be required. Moreover, these units would have to arrive almost en masse to restore cohesion to the defender's posture. Their arrival separately might merely result in each division being subjected to attack by larger enemy forces and thus to defeat in detail.

Once the defender's posture has been physically broken apart and the battle shifted to the rear, mathematical analysis suggests that a decisive victory by the attacker is at least theoretically possible. This especially could be the case if a locally larger attacker is able to isolate the defender's forces in small groups and attack them sequentially. Table 26 illustrates this phenomenon by displaying how many of the attacker's forces are likely to survive a fight to the finish after major breakthroughs have occurred and a maneuver war ensues in the rear areas, assuming all of Army Group A's forces are destroyed. It employs the Lanchester Square equation.

In all four cases the attacker, by breaking through the defender's forward positions, is mathematically able to parlay his original 1.6:1 edge against Army Group A into a decisive victory. Army Group A is completely destroyed (or captured) while large forces of the attacker still survive. The table also has important implications for what could be expected in the next stage of the war if the attacker's surviving forces then advanced against Army Group B. Where the attacker is compelled to engage the defender en masse, behind hasty positions, only 38 percent of the attacker's forces would survive. Even with the addition of the attacker's six divisions already deployed opposite Army Group B, the resulting balance would be 1:1. Army Group B thus would stand a fighting chance of checkmating or even defeating the attacker in a subsequent battle. In the other three cases, however, 64–71 percent of the attacker's forces would survive, and the attacker would gain a force ratio advantage over Army Group B of 1.4–1.5:1, thereby being in a position to defeat it as well in a subsequent maneuver battle.

These data, of course, are illustrative and are highly sensitive to variations in multiple contributing variables. In an actual war, these extreme outcomes would by no means be inevitable or even probable. Nonetheless, these data do illustrate how even at parity, an attacker who intelligently manipulates the dynamics of war can translate a strategy aimed at capitalizing on the defender's weaknesses into a decisive victory. They also suggest how a linear defense, if implemented rigidly and unresponsively, can help pave the way to the defender's defeat.

The issue, therefore, is one of plausibility. Is this scenario realistic—a decisive victory by the attacker and a catastrophic loss by the defender? Or is it a product of a misleading operational and mathematical analysis, one that can be confidently rejected because it is too

Table 26

Range of Outcomes of a Maneuver Battle (percent attacker's forces surviving)

Defender's Formation	Defender in Hasty Positions[a]	Meeting Engagements
Fighting en masse	38	62
Attacked in detail	64	71

[a]Defender units assumed to have lethality twice that of the attacker's units.

implausible to be imaginable? Is a linear array, especially one implemented at low force levels, vulnerable to defeat because the defender is unable to redeploy forces fast enough to compensate for the attacker's ability to be the first to concentrate? If the answer to this question is itself a variable (sometimes yes, sometimes no), under what conditions can the operational advantage swing to the attacker or the defender? And to what extent can the defender's own force size help swing the outcome one way or the other?

LINEAR DEFENSE AS A FUNCTION OF ALTERNATIVE DEFENDER FORCE LEVELS

The defender's force size at parity, relative to the terrain being covered and the number of reserves potentially available in a timely fashion, is a particularly important maneuver strategy. Could a concentrated Soviet attack, in fact, breach the frontal wall that NATO could construct in Central Europe even if no reserves were available? NATO's linear defense would be vulnerable only if the Soviets could gain a major breakthrough that would permit them to advance into the rear areas. Otherwise, they would be unable to take advantage of NATO's lack of reserves. This issue therefore is critically important in evaluating linear defense's future prospects as NATO's forces are reduced.

Although frontal walls in World War II proved vulnerable to rupture, military conditions in Central Europe today differ noticeably from what prevailed then. Some of the changes that have taken place work in favor of the defense and others enhance the offense. However, it is difficult to judge their net effect. Also, any attempt to assess the forward battle today should be tempered by a sense of caution. Past efforts to gauge the offensive-defensive interaction have been notoriously wrong. Before World War I, prognosticators believed that the offense would dominate the next war; 20 years later, they were predicting that the pendulum had swung over to the defense. Their errors in both cases are a sober reminder that military forecasting is at best an uncertain art.

One common contention is that even NATO's present forces would be spread so dangerously thin along the Central Front that they would be easily breached by almost any concentrated Soviet assault. Contributing to this view is the fact that NATO will be able to field only about 45 DEs, a level that compares unfavorably with the 65 divisions that the allies fielded during the 1944 Battle of the Bulge. Impressions, however, can be misleading. Divisions are somewhat larger and much more heavily armed today than they were then. NATO's posture today would have 80 percent of the combat manpower, 90 percent of the artillery tubes, and 60 percent more tanks. It is not dramatically smaller than the posture that Eisenhower had under his command when the Germans attacked in the Ardennes.

Equally important, ground forces have spread out on the battlefield since then, primarily because modern weapons are more mobile and longer ranged than their 1944 counterparts. As a result, NATO's divisions today can defend about 25 kilometers of frontage in a linear array, as opposed to only about 10–12 kilometers for U.S. Army divisions then. NATO's posture today can cover about 40 percent more terrain than the 1944 force. The 1944 posture was able to establish a solid frontal wall along most of its front line, but it had few reserves left over. This is exactly the situation in which Eisenhower found himself in 1944, and it is a primary reason why the allies were fortunate to emerge intact from the Bat-

tle of the Bulge. By comparison, today's posture could defend the same frontage at modern doctrinal standards and still have 15 divisions left over to form reserves.[11]

Since density is relative to the standards of the times, NATO's posture today would not be spread precariously thin as long as it has had the benefit of a month or more of mobilization. Its surprisingly high force-to-space ratio does not mean that today's posture is impregnable or that it could safely absorb future large cuts and still confidently perform a linear defense. But this ratio does suggest that NATO's front wall would be less easily breached than surface appearances suggest and that minor reductions in NATO's posture are unlikely to render it porous.

What dynamics would drive a battle in which a Soviet force, attacking at parity, tried to breach NATO's front wall? One contributing factor would be that weapons today are more lethal, they generate larger volumes of fire, and they move faster than in World War II. Given this explosion of firepower, a future breakthrough battle is likely to be more intense than the Battle of the Bulge, with higher consumption rates and greater casualties. Other things being equal, an intense opening phase of war normally works in the attacker's favor because it speeds up the battle, enabling the offense to exploit the initiative by exerting more pressure on the defense in a short time. In this sense, the dynamics of the modern battlefield seem to work in the Soviets' favor and against NATO.

The nature of the interplay between the offense and the defense, however, probably would affect the outcome more than the tempo of battle. In an effort to gauge how modern weapons might affect this interplay, one thesis holds that the pendulum has swung to the NATO defense in ways that might make the armored battlefield of the future akin to World War I's slaughterhouse. The primary rationale for this thesis is that modern tanks, artillery, attack helicopters, and antitank weapons enable even a small NATO defender to inflict savage losses on a large Soviet attacker as his forces advance over the open battlefield. The implication is that modern battles will be characterized by highly favorable exchange ratios favoring NATO, denying the Soviets the capacity to punch through NATO's frontal wall.

The counter-thesis is that the evolution of modern weaponry has not been unidimensional. The Soviet attacker has benefitted in ways of his own. For example, modern artillery presumably would enable the Soviets to suppress NATO's defense positions. Soviet infantry, in turn, could advance in armored fighting vehicles, protected from NATO's artillery and small-arms fire. Soviet tanks, this thesis holds, could attack in the confidence that their highly sophisticated armor plating renders them invulnerable to all but direct hits from high-speed, kinetic energy NATO tank rounds. Presumably left out in the cold are such NATO defensive staples as shaped-charge tank projectiles and antiarmor weapons that rely on explosive power rather than speed to penetrate. The Soviets also would be able to establish an air defense umbrella that would help keep NATO's attack helicopters and fighter-bombers away from its advancing columns. All of these advantages would permit a Soviet combined arms attack to penetrate NATO's forward defenses in a brief period.

Where does the truth lie between these two quite different theses? In all likelihood, much would depend upon the tactics of the Soviet attack and that of the NATO defense. A typical situation that could arise as part of a larger Soviet theater campaign aimed at concentrating forces at selected points to gain a breakthrough is an engagement in which a single, frontally deployed NATO division is attacked by two Soviet divisions advancing abreast, backed up by a third Soviet division in echeloned reserve. Figure 40 illustrates how these forces would be arrayed.

---

[11]See Department of Defense, 1984.

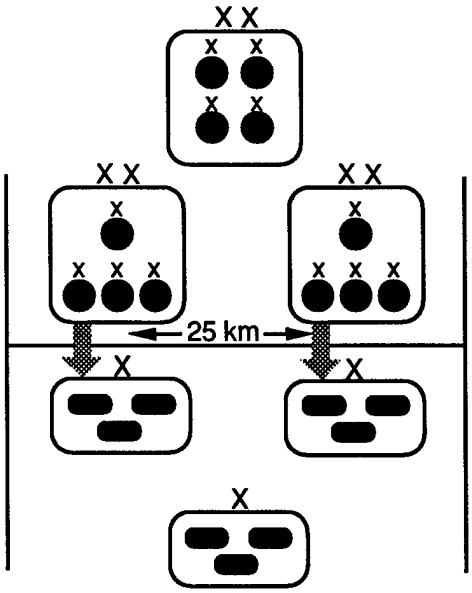

Fig. 40—A Typical Breakthrough Battle

Depending upon its exact configuration, the NATO division would field about 250–300 tanks and a similar number of infantry fighting vehicles with antitank missiles. It also would have some 72 artillery pieces (with another 24 tubes provided by corps support), 60 mortars, 15–20 attack helicopters, and about 2000 infantry troops. Each Soviet division would have a similar number of tanks, infantry fighting vehicles, and attack helicopters, with more artillery tubes but fewer infantry. The balance in maneuver units in this local engagement (DEs) would be 3:1. Since NATO's divisions are larger than their Soviet counterparts, the manpower balance (in DEMs) would be only about 2:1. But since the weapon inventories of the opposing divisions are comparable, the Soviets would enjoy a firepower edge (in WEI/WUV scores) of about 3:1.

Tilting the balance closer to NATO would be the help provided by about 50–100 NATO close air support sorties that could be made available on a daily basis to ground divisions under heavy attack. These NATO air sorties could deliver about 125–250 tons of ordnance each day, or the equivalent in lighter but lethal antiarmor munitions. This would bolster NATO's combined air-ground firepower by an appreciable amount. The Soviets would receive help from their own air forces, but not nearly to this degree.

We will assume that both ground forces are deployed on standard, doctrinally derived frontages for modern military operations. Each Soviet division therefore would be deployed on a 12.5 kilometer front, with two of its regiments initially committed and the other two in echeloned reserve. The NATO division would be deployed on a 25 kilometer front, with two brigades committed on line and the other in reserve. Within each forward brigade, two battalions would be frontally deployed with the third held in reserve. Thus, only four of NATO's nine battalions would be deployed on the front line, enough to establish strength at key locations and continuous line-of-sight contact across the entire division front. The remaining five battalions would be held in reserve, adding depth and staying power.

Popular impressions to the contrary, tactical combat between these two forces would involve a good deal more than the competitive, mechanistic application of firepower. A great deal of movement would occur on the battlefield, and the outcome would be influenced by relative mobility, response time, and force commitments on both sides. Modern Soviet offensive tactics call for the coordinated application of combined arms forces. In the initial engagement, artillery first would attempt to suppress the defender at least temporarily. Immediately thereafter, dismounted infantry forces would try to penetrate the defender's positions and destroy his antiarmor weapons. Following this, tanks would advance quickly, destroy the defender's tanks, and then penetrate several kilometers to the rear.

A succession of these attacks might be launched, perhaps at the rate of two per day on each of three separate attack axes. The Soviets would endeavor to maintain this pace until a hole was opened in NATO's line. Armored and mechanized units in echeloned reserve then would advance forward rapidly, pass through the attacker's outer perimeter, and push out as deep as prudence permits. Simultaneously, other Soviet forces from divisions held further to the rear would transit down the corridor and break out into the open. In this leapfrog fashion, a Soviet attacker presumably would attempt to breach what otherwise would be a formidable NATO defensive wall of firepower.

In response, NATO's defensive tactics would begin with direct and indirect fire to suppress enemy artillery through counterbattery barrages, to drive back enemy infantry, and to destroy attacking tanks as they approached the defense positions. If the effort to rebuff the attack failed, defending NATO units might fall back before the enemy could close on them and reestablish a second defense position a few kilometers to the rear. Meanwhile, NATO reserve forces would converge on the scene in an effort to shore up the defense and establish a more favorable balance of forces.

In all likelihood, the NATO division would use a combination of positional defenses and counterattack maneuvers. The former would aim to channel and then block Soviet advances, leaving enemy forces massed in killing zones that could be deluged with deadly combined-arms fire. The latter would aim to strike enemy columns on their flanks, separating enemy armored units, mechanized units, and artillery in ways permitting their quick destruction. Firepower would function as one important part of a much larger defensive scheme.

The battle between these two opponents would unfold in a complex sequence as Soviet forces tried to penetrate through the depths of the NATO division's zone, stretching some 25 kilometers into the rear. The Soviets would employ mass and firepower as part of their offensive scheme, but their aim would not necessarily be to bludgeon NATO's forces into submission. Rather, they would try to pry apart the defense by suppressing some units while driving others out of position, thereby opening up roads for a quick and deep penetration. NATO's efforts, in turn, would concentrate on keeping avenues of advance blocked as defending forces maneuvered rapidly to new positions and to counterattack. NATO's success would depend heavily upon whether its forces could continue performing this complicated juggling act throughout the course of the battle.

What would be NATO's prospects for averting a breakthrough? Provided it benefits from prepared positions through the division's depth, the NATO division might be able to contain this attack, especially if supporting air strikes were effective. Central to this conclusion is the proposition that by extracting favorable exchange ratios, the NATO division could fight the Soviets to an attrition standstill. Several risk factors could work together to unhinge this NATO division, however, especially if it was not reinforced promptly from outside.

A particular imponderable is that inexperienced soldiers sometimes lose their poise when first exposed to the shock of an enemy attack. Panic can ensue, commanders can lose control, and units at key positions can lose their ability to function as a cohesive team. Something like this happened to French forces at Sedan in 1940, to the Russians at Barbarossa, and evidently to some American units at the Battle of the Bulge. The result can be a very quick enemy breakthrough, conceivably within the first few hours of fighting and long before either side has suffered major attrition.

Another risk factor is that NATO's commanders might make tactical mistakes at inopportune moments. A classic example is the American failure at the Battle of the Bulge to coordinate the withdrawals of the 14th Cavalry Group and the 106th Infantry Division, which opened a gaping hole in the defense line. Another example is the French maladroit coordination of armored counterattacks at Sedan in 1940. Mistakes like these are an inevitable part of combat, but they could have an especially disastrous effect on the contemporary battlefield because modern armored/mechanized forces can move so rapidly to exploit them.

An unanticipated breakdown might occur in the weapons of one part of the defender's combined arms team. Many modern weapons are untested in battle; major failures occasionally have occurred in the past. Allied antitank missiles could fail to penetrate Soviet tanks. The consequence could be a profound change in the dynamics of unit engagements that, while correctable over time, might have an early deleterious effect.

The defending force might become spread too thin to continue blocking the main axes of advance. Although a NATO division initially has sufficient forces both to cover its frontage and to withhold sizable reserves, casualties during battle can eat away at this margin of safety. High daily loss rates, initially of up to 25 percent for armored vehicles in main sectors, are expected in modern combat. Within a few days, a NATO division could find itself reduced by 50 percent in strength. Moreover, it might well have been driven backward beyond the rear perimeter of its prepared positions and therefore denied any further protection from them. Concurrently, local enemy penetrations might have expanded that division's frontage by 50 percent. Especially if its units had been physically separated, the division's capacity to continue covering its assigned sector in enough density could be dangerously eroded. The result could be a breakthrough within a few days, through the combined mechanics of attrition and front line expansion.

None of these risk factors means that NATO's defense line necessarily would buckle at the point of impact. The Soviets would face parallel vulnerabilities of their own, plus some that are uniquely theirs, a rigid command and control system among them. The Soviets did become skillful at fighting breakthrough battles in World War II, and prudent NATO planning can hardly afford to risk that they have not carried these skills forward. Were they to attack with similar effectiveness today, the interplay of these risk factors could create unmanageable pressures on NATO's forces.

Statistical dynamics enter this breakthrough calculus when it is applied across NATO's front wall. A single division might not fail badly enough to allow large enemy forces to pass through, but if several divisions were attacked in strength at once, the combined probability that one of them would crumple could be high. For example, suppose that three NATO divisions lined up abreast are attacked concurrently. If the odds are only 30 percent that any one of them might break, the combined probability that at least one of them would fail would be 66 percent. A 50 percent chance of any one division failing would translate into nearly a 90 percent chance of a breakthrough occurring somewhere.

The gaining of a breakthrough against one division could enable enemy forces to assault the other two divisions from the flanks and rear, opening a large hole in the defender's line. Considerations such as these helped enable German commanders to confidently pursue breakthroughs in World War II. They are a primary reason why prudently minded planners today often assume that despite the proliferation of defensive weaponry over the past two decades, a NATO frontal wall still would probably be breached.

Static Planning Factors as a Methodology

As the Battle of the Bulge showed, an enemy breakthrough does not necessarily spell defeat for the defense, even one deployed on a linear array. Much depends on whether the defense can respond in time by redeploying reserves either to contain the breakthrough or to compete effectively in the subsequent war of maneuver in the rear areas. It is here that the variables of defensive force levels and theater employment doctrine enter the modern military equation. The issue of whether, and at what point, NATO no longer will be able to rely on a linear array largely turns on its ability to respond in this fashion even as its forces in Central Europe are reduced.

Assuming parity, does the offense-defense interaction depend on the absolute level of forces on the defender's side? In the face of a concentrated enemy attack, would NATO be able to respond to breakthroughs only as long as its forces are maintained at their present levels? Would NATO's capacity to respond from a linear array diminish as its forces are reduced, even though Soviet forces are themselves reduced by the same margin? If the phenomenon of diminishing capacity does occur, how serious is it? In particular, at what point of lower force levels would NATO be compelled to abandon its frontal wall and adopt a new doctrine, one that would enable it to respond better?

Before 1989, little analytical attention had been devoted to developing a theory of NATO force requirements for a linear defense at parity and low force levels. As a result, initial attempts to grapple with this subject relied on rule-of-thumb static planning factors to help provide a sense of force needs. These factors tended to generate a high theory of requirements. A 1988 German White Paper, for example, postulated that 48 divisions were needed to defend Central Europe.[12] That would provide enough reserves to make NATO's posture something other than a linear defense as defined here, somewhere between a linear and an echeloned defense.

These planning factors stemmed from a military methodology that sized NATO's forces according to the Central Region terrain they were defending, not according to the enemy force they were fighting. This methodology was anchored on a linear employment doctrine that sought to deemphasize reliance on forces either deployed at the far corners of the theater or otherwise unlikely to be available in responding to breakthroughs. The operative goal of this methodology was to generate readily available NATO forces to cover the terrain in sufficient breadth and depth to deny Soviet forces any scope for maneuver. The methodology therefore called for a strong front wall that could not easily be breached, backed up by the large reserves that military commanders traditionally have regarded as desirable. It arrived at a theory of requirements for linear defense through a four-step process.[13]

---

[12]FRG Ministry of Defense, 1988.

[13]Department of Defense, 1984.

- It calculated the number of divisions needed to populate the inter-German border on doctrinally derived frontages of about 25–30 kilometers apiece. With a Central Front of 750 kilometers, this created a requirement for 24–30 divisions to man NATO's front line.[14]
- Each of NATO's eight forward corps sectors was to be provided a covering force of an armored cavalry regiment to conduct screening and reconnaissance missions along the border. This force equated to 2-2/3 DEs.
- The methodology called for each forward corps sector to be provided a backup reserve division, creating a requirement for eight additional divisions. The idea here was to guard against what occurred at the Battle of the Bulge: an early penetration caused by tactical mistakes or the shock-induced unravelling of some units. Another purpose was to give each corps sector additional staying power against a concentrated Soviet attack. An additional division, illustratively, might be expected to elongate the staying power of each corps from three to seven days, providing added time for reinforcements to converge at the scene.
- Sizable reserve forces were to be assigned to the NORTHAG, CENTAG, and AFCENT commanders. Assuming a three-division corps for each command echelon, this added nine divisions to the NATO total.

The four building blocks of this methodology, when added together, created a requirement for roughly 40–50 NATO divisions, as displayed in Table 27. The policy implications of this estimate are apparent: Since NATO's posture currently fields only 45 divisions at M+30, fairly small reductions (beyond 10 percent) would unhinge its defense.

This methodology provides one conceptual lens for viewing NATO's force needs for a linear defense. Its strength lies in its capacity to provide a theory of military stability at parity by providing both sides a highly assured defense posture. With 40–50 divisions dominat-

Table 27

NATO Ground Force Needs According to Traditional Planning Factors

Force	Low	High
Covering force	2-2/3	2-2/3
Divisions on line	24	30
Corps reserves	8	8
Theater reserves	9	9
Total	41-2/3	49-2/3

[14]The issue of divisional frontages has been a controversial one. The concept of average frontages of 25–30 kilometers has been a staple of NATO planning for two decades but has begun to change recently as more mobile weapons and some flexible concepts have entered Western military thinking. A standard U.S. or German division can cover up to 50 kilometers and still maintain constant line-of-sight coverage on normal rolling terrain. Stretching a division's frontage that far, however, would cause all of its reserves to be placed on the frontline, thereby sacrificing its sustainability. As a result, a tradeoff arises between breadth and depth; depth can be sacrificed only if ample reserves from other areas can be brought to bear to contain breakthroughs. The critical issue shaping frontages thus is sustainability goals. If a 25–30 km frontage can provide three days of staying power, is this enough, or would less suffice? If so, longer frontages are acceptable. For a thorough analysis of related issues, see Paul Davis, "Prospects for Military Stability in a Deep-Cuts Regime," in Cuthbertson and Volten, 1990.

ing the Central Region terrain on each side, neither opponent would have much confidence in succeeding in an attack against the other.

The drawback of this methodology is that it provides an ideal standard, free of resource constraints, which gives no basis for gauging the effect of force reductions beyond minimal amounts. Whether this methodology provides a technically confident estimate of minimum force thresholds for linear defense is uncertain. No estimate is made of how far the Alliance's confidence would decline if NATO's forces fell below this minimum threshold. Would a minor deficiency produce a major risk or a minor one? Would a wholesale change in NATO's employment doctrine be needed, or only a modest one?

This methodology's inability to address these questions stems from its reliance on static planning factors that derive from terrain, mission, and staying power estimates. As a result, there are no analytical techniques for dynamically assessing the critical issue in judging linear defense: At what point of reductions does NATO's posture become so dispersed that adequate forces cannot promptly redeploy in response to an enemy breakthrough attack? The methodology implicitly assumes that any posture below 40 divisions would render NATO incapable of responding in a timely fashion, but it does not show exactly why.

By definition, a NATO posture below 40 divisions would have less force available in absolute terms. But so would the Soviets. It is the interaction between the two sides that matters. Even if NATO was constrained from thinning out its present front line, reserves would still be available at all levels above 24 divisions. For example, a posture of 35 divisions would provide 11 reserve divisions. That would be five fewer than the 16 reserve divisions provided by a 40-division posture, but the Soviets themselves would have five fewer divisions with which to attack. Taking into account this relative drawdown on both sides, would NATO be less able to defend at 35 divisions than at 40 divisions, and why? What about 30 divisions? This methodology cannot answer these questions.

Toward an Interactive Analysis: Limiting Cases

A methodology that allows for an interactive analysis provides a better basis for addressing these difficult questions. The idea that a linear defense becomes more vulnerable as the defender's mass decreases is based on the proposition that battle dynamics would increasingly benefit the offense as absolute size declines on both sides. When this proposition is pushed to its limits, taking into account offense-defense interactions, it makes intuitive sense. If both sides had 100 divisions, the terrain would be so dominated by ground forces that neither side would have room to maneuver against the other. But if both sides had only ten divisions, the terrain would be open to movement. If the offense was free to maneuver, and the defense was prohibited from doing so by its own employment doctrine, the offense would benefit, and linear defense would become more difficult as forces on both sides were reduced.

Consider first the high end of the spectrum. Since the Central Region front can accommodate only 60 first-echelon divisions deployed on attack frontages, both sides would be able to form a virtually impregnable front wall while withholding 40 divisions in reserve. The Soviets would be physically unable to concentrate larger forces anywhere directly on the front. They therefore would be compelled to mass their forces in their rear areas and to advance in successive waves of echeloned formations against a defense that itself was echeloned in depth. Since the lead echelons would be compelled to attack a prepared NATO defense with engaged ratios of no more than 1:1, the Soviet advance would proceed very

slowly. Even if NATO's lead echelon buckled, sizable reserves would be located nearby to redress the situation. Meanwhile NATO would have ample time to bring its full 40-division reserve to the main point of contact. The war between them would take the form of an attrition contest with little movement in any direction, much as occurred in World War I.

Now consider the other extreme in which both sides are reduced to ten divisions. Further assume that the Soviets are completely free to concentrate and that NATO is compelled to distribute its forces uniformly across the 750 kilometer front. NATO therefore is left with a porous line and no reserves. The Soviets would be able to descend on one or two NATO divisions, quickly surround them, and destroy them. They could then proceed to the next segment of NATO's forces and sequentially repeat the process until the end was reached. Seeing the first Soviet attack begin, NATO might then be able to start concentrating its forces, but it would face a desperate race to form enough mass before the battle had gotten hopelessly out of hand. A Soviet victory in this case would not be guaranteed, but NATO clearly would be far more likely to lose than in the 100 division case.

These two cases help establish the proposition that while parity may be a necessary condition for stability, it is not a sufficient condition when the attacker is able to concentrate faster than a defender inhibited by his own employment doctrine. Absolute mass enters the military equation. The larger the defender's posture relative to the terrain being protected in a linear fashion, the better able is the defender to react in a timely fashion by bringing adequate reserves to the main point of contact.

Interactive Analysis: A Preliminary Examination

These two extreme cases alone are insufficient to permit analysis of how the risks shift along the spectrum of force levels being considered in today's policy arena. It is necessary to examine how force reductions would affect the operational capabilities of both the Soviets and NATO.

For many years, the Soviets were assessed as being capable of marshalling a 90–100 division posture in Central Europe, or enough units to create five "fronts" of about 20 divisions apiece (Fig. 41). By Soviet doctrinal standards for offensive operations, a single front typically would deploy along a 150–350 kilometer sector. As a result, the Soviets were deemed capable of covering the Central Region's entire 750 kilometer border with three or four "fronts." Meanwhile, one or two fronts could have been withheld as a second strategic echelon that could be committed, if necessary, to sustain an attack.[15]

Even without this second echelon, a three front attack would have provided the Soviets impressive offensive options, especially for fronts A and B in the example displayed in Figure 42. For each front, one army composed of four divisions might have concentrated on a 50 kilometer subsector, with a second army adjoining it on a 50 kilometer subsector. A third army might have deployed more thinly on the remaining 100 kilometers, enabling each front to launch two concentrated attacks, while still withholding two full armies as a tactical reserve. Meanwhile Front C would have been able to launch secondary attacks aimed at pressuring NATO's forces there and pinning them down. If necessary, front C could have been reinforced by 5–10 divisions from the second strategic echelon to increase this pressure.[16]

---

[15]See Erickson et al., 1986. See also Department of the Army, 1978.

[16]Department of the Army, 1978.

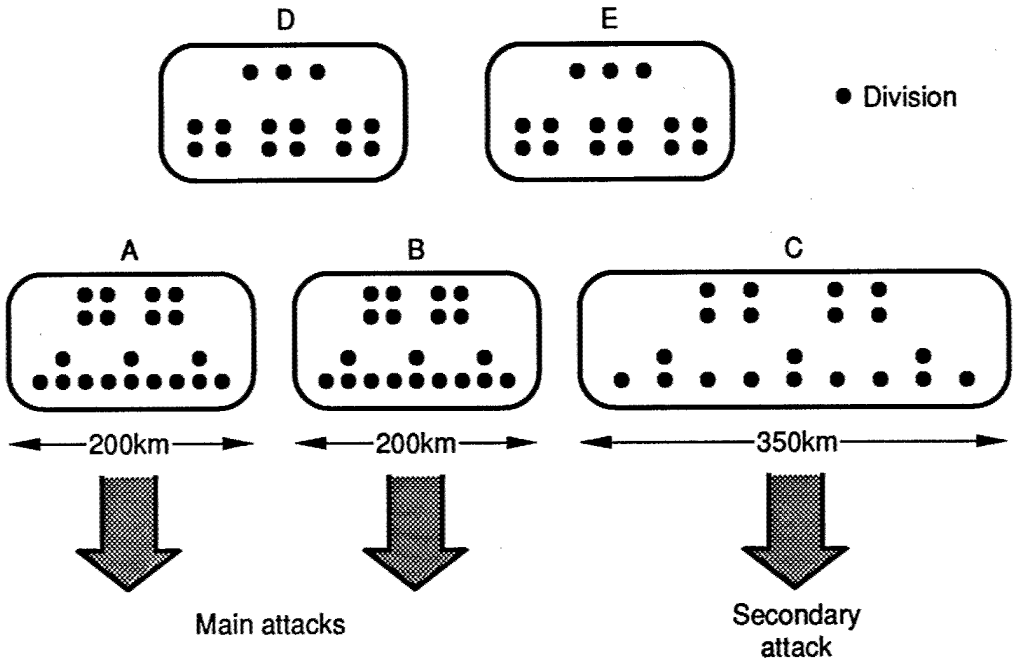

Fig. 41—Traditional Soviet/Warsaw Pact Five-Front Attack Formation

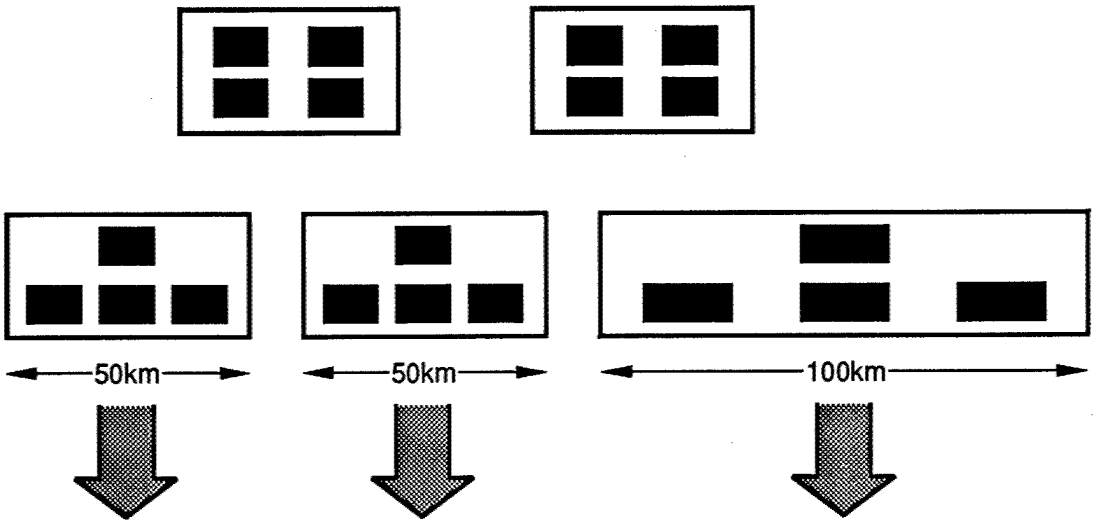

Fig. 42—Soviet Front Attack Formation (200 km Sector)

By working in tandem, these three fronts would have enabled the Soviets to launch concentrated assaults on two 100 kilometer segments of NATO's defense line, while maintaining some pressure on the remaining 550 kilometers. If these two attacks failed to gain a major

breakthrough, forces from the remaining second-echelon fronts could have been brought forward to continue the operation. Eventually a major breakthrough would have been gained, allowing large Soviet reserves to pour through the gap. It was this kind of offensive pressure—concentrated at points but sustainable and distributed broadly—that lay behind NATO's traditional concerns for its forward defense capability.

The operational effect of the Soviet/Warsaw Pact force reductions now underway must be assessed against this historical background. As noted earlier, a CFE I Treaty will leave the Soviets with up to 65 divisions that could be marshalled against NATO in Central Europe. This force would deprive the Soviets of the two fronts that have provided their second strategic echelon, but the remaining force still will be large enough to form three fronts capable of extending doctrinally planned pressure across nearly all of NATO's linear array.

Lower Soviet force levels (e.g., 24–48 divisions) might be achieved as a result of a CFE II accord, reducing the Soviets to well below their requirements for a full three-front attack. Since the Soviets would be left with one or two reinforced fronts, they would not be entirely deprived of offensive options, but they would be compelled to make major changes in their own theater employment doctrine that would reduce their ability to exert pressure along NATO's entire layer cake. Reduction to a 48-division posture would have the least onerous effect, but as further cuts were made, the constraints would mount. Table 28 displays how Soviet forces might be distributed at each force level.

The Soviets would probably choose to deploy as large a force as possible in their primary sector to execute an offensive strategy against NATO's linear array there, compelling them to reduce their forces in their secondary sector to only a reinforced army. That would not be enough to launch even a secondary attack, but it would establish a screening force along the border region and achieve a .5:1 ratio against NATO forces committed along the front line.

Even with this stingy allocation to their secondary sector, Soviet offensive options in the primary sector would decline steadily as force levels diminish. In the case of 48 divisions apiece, the Soviets would be able to cover the NATO's Army Group A[17] line with two full-sized fronts, each deployed on a standard attack frontage. As a result, they would be able to

Table 28

Soviet Theater Employment Strategy at Low Force Levels: Base Case

	Primary Sector		Secondary Sector	
Divisions	Size (km)	Divisions	Size (km)	Divisions
60	375	42	375	18
48	375	42	375	6
42	375	36	375	6
36	375	30	375	6
30	375	24	375	6
24	375	18	375	6

[17]NATO's Army Groups currently are named NORTHAG and CENTAG. This command structure probably will change in the years ahead, so this analysis will postulate that NATO will retain two Army Groups, renamed as A (on NATO's left) and B (on NATO's right). This nomenclature is in keeping with the general model developed above.

launch four concentrated attacks, each 50 kilometers wide, in search of a breakthrough. Alternatively, they could amass much larger forces in support of one or two attacks. At the other extreme of only 24 total divisions apiece, they would be compelled to advance against NATO's Army Group A with only one front (minus two divisions). This force would still be capable of launching one or two concentrated attacks, but a larger portion of forces (e.g., 33 percent vs. only 15 percent) would have to be deployed to pin down Army Group A's forces elsewhere along the front line. The Soviets would then be able to generate only 12 divisions as a breakthrough and exploitation force, nowhere near the 36 divisions that the two-front posture could provide.

Table 29 displays how this Soviet theater employment strategy would interact with NATO's linear defense at these different force levels. Also shown is an alternative Soviet strategy that deploys enough additional forces opposite Army Group B to maintain a .5:1 ratio, not only against NATO's frontline forces there but also its corps and Army Group reserves. The table assumes that NATO's forces at each level are evenly divided between the two sectors; redeployments will be examined below.

On surface appearances the table suggests that, if anything, the phenomenon of declining force levels at parity operates against the Soviet attacker and in favor of the NATO defender. A 1.50:1 ratio is one commonly used threshold for gauging the attacker's success in an operation across a wide, multicorps sector of Army Group size. If the Soviets were to pursue the alternative strategy, they would never exceed this threshold; the ratio stays constant as force levels decline. By pursuing the more aggressive base case strategy, the Soviets would be capable of amassing a 1.75:1 advantage at 48 divisions apiece. However, their edge would decline steadily as force levels on both sides were reduced, because at each successive stage the Soviets lose double the number of divisions against Army Group A as NATO loses (six vs. three), moving the force ratio downward. Against Army Group B, meanwhile, NATO loses three reserve divisions at each step and the Soviets lose none. But since Soviet forces are small, NATO incurs no danger. By the time the 24 division level is reached, the Soviet advantage against Army Group A is reduced to the 1.5:1 threshold, a much more manageable proposition for NATO than the 1.75:1 disadvantage at the highest force level.

Further reinforcing the conclusion that lower force levels work to the defender's advantage is the fact that NATO's dependence on outside reinforcement (from Army Group B or elsewhere) declines steadily as forces are reduced. In the 48 division case, NATO requires four additional divisions to achieve a 1.5:1 ratio. At each step downward, one division is

Table 29

Soviet/NATO Force Ratios

	Base Case Soviet Employment Strategy		Alternative Soviet Strategy	
Force Levels	Primary Sector	Secondary Sector	Primary Sector	Secondary Sector
48	1.75	.25	1.50	.50
42	1.71	.29	1.50	.50
36	1.66	.33	1.50	.50
30	1.60	.40	1.50	.50
24	1.50	.50	1.50	.50

subtracted from this total. By the time the final stage is reached, NATO's NORTHAG forces seem adequate without any outside reinforcement. For this reason, the viability of linear defense—calling for little reliance on force redeployments from one Army Group to the next—seems to increase as forces decline.

These clearly are powerful arguments that not only suggest deep force cuts may be stabilizing but also raise questions about whether there would be any need to alter NATO's linear defense doctrine. Appearances, however, can be deceiving. Powerful arguments to the contrary require an examination of how Army Group A's defense effort would encounter growing constraints of its own as NATO's force levels are reduced.

The geometry of the Central Region battlefield complicates NATO's defense task. A straight line drawn between two end points will clearly be shorter than a concave line. A larger amount of NATO's force is required to defend its concave line at all points with comparable density and counter-concentration capacity. NATO's concave line demands some 15–20 percent more divisions, frontline and reserve together, than a straight line. This adverse geometry is one factor inhibiting deep NATO force cuts.

Even with a concave line, deep force cuts at parity, in principle, would pose no special dilemma as long as the defender is not obligated to meet the main point of an enemy attack directly on the border. The defender simply could mass his forces at a central location and march them forward to meet the attacker in a meeting engagement. With a straight line—assuming the attacker is moving from a similar central location and the defender moves with equal speed—the battle would be fought very near the border regardless of where the attacker chose to cross. With a concave line, the attacker could reach the flanks of the border faster than the defender because he would have less ground to cover. He would then be able to advance some distance into the defender's territory before the defender's forces could arrive in strength to bar his advance and give battle. But within a few days, the defender would be capable of giving battle on equal terms.

Political considerations have made this military tactic impossible for the Alliance. NATO has been compelled to deploy strong forces in the forward areas to ensure that a Soviet advance could be met directly on the border. Further, uncertainty about exactly where the Soviets might cross, coupled with their capacity to reach many locations faster than NATO forces could arrive and dig in, has led NATO to establish a strong line along the entire border. This requirement has pulled NATO's center of gravity well forward, resulting in a large portion of NATO's early posture being deployed directly on the front line.

It is here that this political constraint on abandoning NATO's linear array has entered to cloud the force reduction picture. Without a decision to abandon this aspect of a linear array, deep force cuts almost inevitably would result in NATO's reserve forces being sacrificed first and its frontline units last. Taken far enough (50 percent cuts), NATO could be left with a strong front line but with all its forces scattered along the border and supported by no reserves. This deployment, in turn, would leave NATO's posture vulnerable to a concentrated Soviet attack somewhere near the center, aimed at enveloping either side of NATO's line.

NATO's forces deployed away from the main point of attack could begin counter-concentrating almost immediately, but relative motion dynamics would work against NATO. Because of NATO's concave line and longer distances, Soviet forces could begin giving battle before all of NATO's forces could arrive. Moreover, advancing Soviet forces would exert leverage over NATO's redeployment pattern. NATO would be uncertain exactly where the attack was coming, so only the forces deployed at the far periphery could begin moving immediately, with the rest redeploying in a ripple fashion converging on the center. Since

these forces would have the farthest distance to travel and their neighbors would begin moving somewhat later, NATO's buildup rate in the battle area could take several days to complete. The risk is that the Soviets might be able to gain a major breakthrough by then.

The movement of NATO's defense to Germany's new eastern borders will alter this geometry by eliminating NATO's old concave line and its exterior lines of communication, as Fig. 43 shows. But it will not produce a straight line or a convex line that is ideal for a linear defense. Instead, it will produce an irregular line with a long salient jutting eastward. To the extent that NATO can focus its defense plans solely on defending eastern Germany's Oder-Neisse line, the task of quickly redeploying NATO's reserves will become much easier. But if NATO must simultaneously defend Bavaria in strength, the task might well become more difficult because even longer distances would have to be covered. This difficulty could be overcome if NATO's defense line were dropped through Czechoslovakia from Germany's eastern border southward. Short of this politically controversial step, NATO will continue to labor under constraints deriving from adverse battlefield geometry.

Both a strong front line and local reserves matter heavily in NATO's defense calculus. Table 30 displays how NATO's forces might be allocated at all five levels, assuming they are

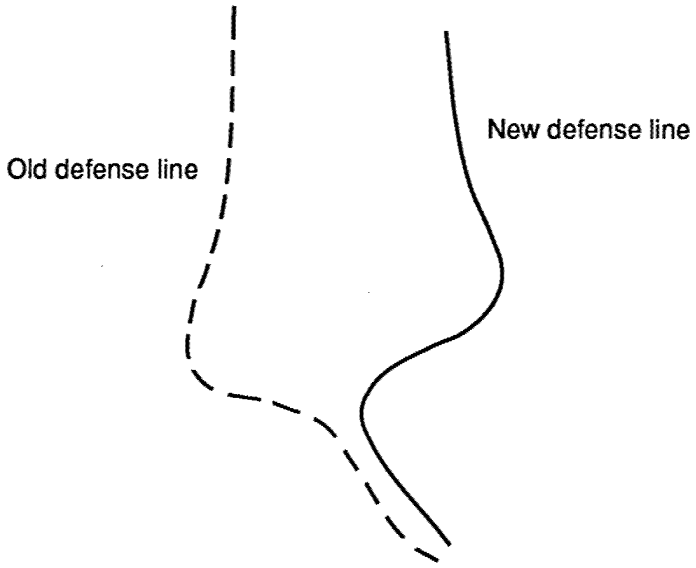

Fig. 43—Geometry of the New Battlefield

Table 30

Illustrative NATO Force Allocations

	Army Group A					Army Group B				
Total theater	48	42	36	30	24	48	42	36	30	24
Total Army Group	24	21	18	15	12	24	21	18	15	12
Front line	12	12	12	12	12	12	12	12	12	12
Corps reserves	6	6	3	3	0	6	6	3	3	0
Army group reserve	6	3	3	0	0	6	3	3	0	0

initially divided equally between Army Group A and Army Group B. A posture of 48 divisions provides Army Group A a strong in-depth defense. Each of Army Group A's four forward corps receives three frontline divisions that are deployed on doctrinally prescribed frontages of about 30 kilometers. Backing up this strong line, each corps also has a 1-1/2 division reserve, counting its armored cavalry regiment, that is immediately available either to its parent formation or to neighboring corps. Army Group A itself has a sizable reserve of six divisions that can come forward quickly. The result is an Army Group that covers the terrain in high density and can perform the full range of missions associated with a defense effort. Breakthroughs are unlikely to occur quickly; if they happen, they stand a good chance of being contained.

Army Group B's forces can support Army Group A if the main attack comes there, and vice versa. Consistent with linear defense planning standards, the analysis assumes here that Army Group B's 12 frontline divisions would need to remain on line. They need to guard against the risk that some large portion of the Soviets' echeloned reserves might divert there; furthermore, pulling them out, under enemy fire, might be difficult and costly. But that would leave a large force of 12 divisions—corps and Army Group B reserves—that could redeploy northward, thereby increasing Army Group A's posture by 50 percent.

Finally, large tactical air forces would be available to help NATO's ground units. Assume that 3800 tactical combat aircraft are available to each side when the ground balance is 48 divisions apiece and that these forces are reduced proportionately as the ground postures are drawn down. With a posture this large, up to 1600 aircraft could be available for close air support and battlefield air interdiction sorties (after discounting for other missions). The Soviets would have a similar air force, but probably fewer aircraft would be used in the ground attack role and they would be less effective. NATO thus would have an air advantage that would add a margin of insurance.

Deep force cuts would eat away at these strengths in absolute terms. An important change is that Army Group A's reserves would be disproportionately reduced: A 25 percent cut in Army Group A's posture yields a 50 percent cut in reserves. In theory, NATO could thin out its front line in tandem. But since this would leave the entire line more vulnerable to a quick breakthrough, let us assume that keeping it strong is the highest priority. Later, the analysis will redeploy some of these forces to help counter a breakthrough, thereby modeling the effects of using them as reserves after the battle has begun.

NATO's reserves thus are sacrificed first. Reduction to 36 total NATO divisions (18 in Army Group A) would drop each corps reserve to three-quarters of a division, reducing the staying power of each corps. Meanwhile, a three-division Army Group A reserve corps would also be lost. Reduction to 24 NATO divisions would leave Army Group A totally bereft of committed reserves. Without tactics that pull units out of the line in less threatened sectors, Army Group A would have no ability to halt a breakthrough, should one occur.

Deep cuts also would erode the number of Army Group B reserves that could be redeployed to Army Group A. Once again, the effect on Army Group A's posture would be disproportionately high. A reduction to 36 total divisions would not only cut NORTHAG's reserves by 50 percent but Army Group B's reserves by 50 percent. The total force available to Army Group A would be cut from 36 divisions to 24. A posture-wide cut of 50 percent (to 24 divisions) would produce a total Army Group A reduction of 66 percent to 12 divisions. Depleted also would be NATO's comparative advantage in tactical air power, especially if reductions were implemented in a fashion that sacrificed ground attack forces to preserve NATO's air defense capability.

Interactive Analysis: An Early Breakthrough Scenario[18]

Both the Soviets and NATO would lose important capabilities, in an absolute sense, as their forces are reduced. The issue therefore becomes whether one side or the other would gain in a relative sense. At higher force levels, the Soviets would be able to launch a powerful attack, but NATO would have a much better posture at its disposal to defend. At lower force levels, the Soviets could attack less strongly, but NATO would not be able to defend as well. The argument that the Alliance's defense prospects would decline as forces are reduced is based on the proposition that NATO's defense effort would become relatively weaker as the cutback process unfolds. If that is true, it could happen either at a uniform rate with a cumulative effect, or all at once in a single stage. But for the proposition to be true, it would have to be a substantial weakening.

Do relative drawdown dynamics, in fact, work against NATO? Figure 44 displays how Soviet and NATO forces might be deployed at a war's onset. For simplicity's sake, the figure shows only a concave line, but the basic physical relationships would not change a great deal if a straight line, or one reflecting NATO's new borders, was portrayed. The figure shows the two extreme cases considered here, 48 and 24 divisions apiece, assuming that the Soviets have gained the initiative by concentrating first, at NATO's center. NATO has not been allowed to undertake any counter-concentrations; its deployments consequently reflect initial force locations for executing a linear defense. Notice that in the 48-division case, the Soviets have large forces (26 divisions) in echeloned reserve behind the main point of attack. NATO has 24 divisions in corps and Army Group reserve, scattered across the battlefield. At each successive stage of reduction, these reserves on both sides would be pared back. The end point would be the 24-division case, where the Soviets would have only eight reserve divisions, but NATO would have none at all.

At all points along the spectrum of force reductions, NATO's defense prospects would hinge on its capacity to counter-concentrate fast enough:

- To prevent a penetration of the front line.
- To contain a penetration before it became a major breakthrough.
- To fight a rear-area maneuver battle once a breakthrough had become uncontrollable.

Table 31 helps analyze NATO's relative capacity to achieve each of these tasks by displaying five critical force ratios. The first ratio shows the overall balance against Army Group A, assuming all but six Soviet divisions are concentrated there. The second ratio reflects all Soviet forces available for commitment to the main sector in relation to NATO forces immediately deployed in that sector. The main sector is assumed to be 75 kilometers wide. Soviet forces include all but 12 divisions, six opposite Army Group B and another six deployed on secondary sectors against NORTHAG. NATO's forces include three frontline divisions, plus whatever corps reserves are immediately available. The engaged force ratio within this main sector assumes NATO's forces there are attacked by a first-echelon Soviet reinforced tank army (six divisions). The fourth force ratio displays all Soviet forces available for commitment to the main sector against all the forces that Army Group A could muster, over time, to

[18]The following analysis of early and later breakthrough scenarios assumes, for analytical purposes, that the defender rigidly adheres to a strict interpretation of linear defense as defined here. It allows for none of the doctrinal adjustments discussed in Sec. 3 that could help alleviate the risks. Thus, the central conclusion is not that defense at low force levels is hopeless, but that a rigidly linear philosophy becomes even more vulnerable as NATO's forces are reduced. Doctrinal innovations could reduce many of the risks outlined here.

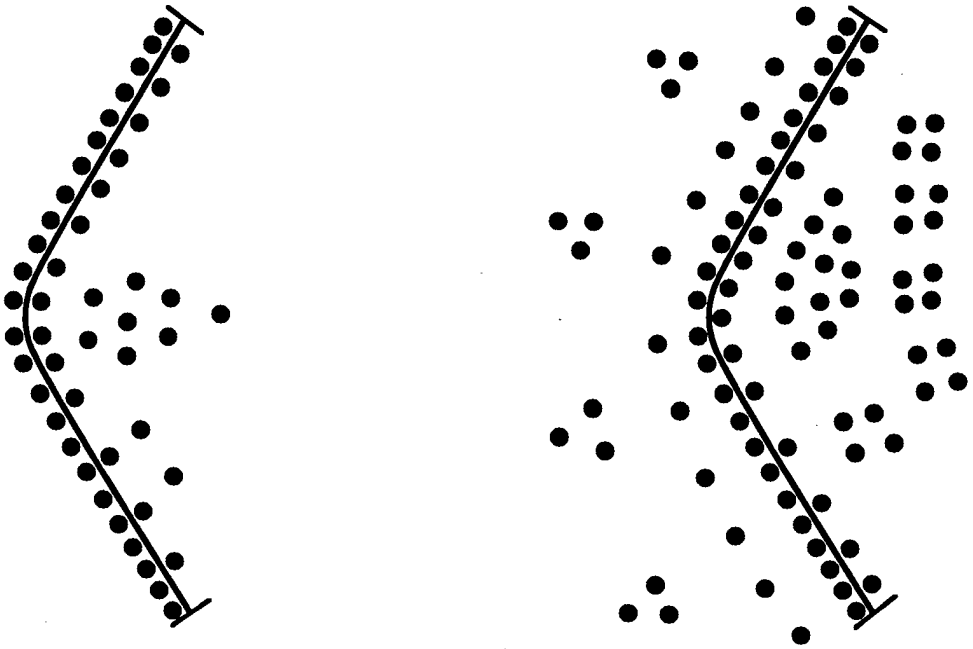

Fig. 44—Illustrative Deployments on D-Day (Divisions)

Table 31

Soviet/NATO Force Ratios for Breakthrough Battle

Ratios	NATO Force Level (divisions)				
	48	42	36	30	24
1. Overall Army Group A	1.75	1.71	1.66	1.60	1.50
2. Main sector	8.0	6.7	6.4	6.0	4.0
3. Engaged	1.3	1.3	1.6	2.0	2.0
4. Breakthrough vs. Army Group A	2.0	2.0	2.0	2.0	2.0
5. Breakthrough vs. Army Groups A and B	1.20	1.25	1.33	1.50	2.0

employ in this sector. It includes not only corps and Army Group A reserves, but also three divisions that are presumed to be pulled out of the line elsewhere and redeployed. The final force ratio displays the same Soviet forces assuming these Army Group A units are reinforced by Army Group B's corps and Army Group reserves. Attrition will be considered below.

Ratios 1 and 2 show that at high force levels, the Soviets could employ their large overall advantage to concentrate against a main sector and achieve very sizable advantages. The maximum is 8.0, but at 42 divisions, the ratio is still nearly 7:1. As a result, the Soviets should be able, NATO redeployments aside, to exert the immediate and sustained pressure

that normally leads to a breakthrough even against a tough defense. At lower levels, this advantage diminishes. Nonetheless, it remains fairly high. Between 30 and 24 divisions, it ranges between 6 and 4:1. This advantage well exceeds the range of 2 or 3:1 that is one threshold for gauging a prepared defense's ability to keep a main sector attack at bay.[19]

The Soviets would probably encounter greater difficulty in gaining a very quick breakthrough at higher forces than at lower levels because the engaged ratio is only 1.30:1 as high force levels, a product of the first-echelon Soviet tank army encountering a NATO three-division frontline force that has 1.5 divisions of corps reserves immediately to the rear. At lower force levels (30 divisions and below), however, this situation changes in the Soviets' favor because NATO's corps reserves have been stripped away. The engaged ratio consequently rises to 2:1.

An engaged 2:1 ratio is theoretically manageable, but only if the defender's frontline divisions perform effectively and conduct a coherent effort that transforms the battle into an attrition contest. An entirely different situation would come to pass if one or more of these divisions were to suffer the fate of the French at Sedan or that of the American forces at the Battle of the Bulge. By this is meant either an internal unravelling from the shock of intense combat or tactical errors that open holes in the line leading to encirclements in the forward areas. Both of these setbacks are not uncommon for troops and commanders lacking combat experience. If either or both were to occur, the consequences for NATO at low force levels could be serious. NATO's forces in the main sector would completely lack corps reserves, thus exposing their rear areas to a rapid enemy advance if a breakthrough does occur.

NATO's posture at higher levels would seem to be less vulnerable to this kind of setback precisely because some reserves would be available. However, the Soviets plausibly could compensate for the reduced likelihood of a shock-caused breakthrough at any single point by employing their large forces to expand the width of their main attack. This would bring more NATO divisions under intense pressure, albeit all benefitting from reserves. For analytical purposes, this analysis will assume that a quick breakthrough does, in fact, occur at all five force levels.

Ratios 4 and 5 help perform this task. Ratio 4 shows that at all five levels, the ratio after Army Group A concentrates all its potentially available forces is a constant 2:1. This seems puzzling since the Soviet overall advantage of 1.75:1 at the highest level diminishes to 1.50:1 as the forces on both sides are reduced. The explanation, however, is simple. At low force levels, NATO is able to redeploy only one-third of its forces that initially are located outside the main sector. As a result, the Soviets gain an advantage in the main sector that exceeds their overall edge in Army Group A (2.0:1 vs. 1.50:1). As force levels rise NATO is able to redeploy about 70 percent of its forces located elsewhere. The Soviet advantage in the main sector still exceeds their overall Army Group A edge (2.0:1 vs. 1.75:1) but by a smaller amount than at lower force levels. The result is a constant 2.0:1 ratio in the main sector after Army Group A redeploys.

While a 2:1 ratio might be manageable when the defense is fighting from prepared positions on a cohesive front, it would be an undesirable state of affairs if the Soviets were able to break through NATO's line to the point of gaining unimpeded access to the rear areas. In that circumstance, NATO could find itself deprived of nearly all the classical advantages of fighting on the defensive, dramatically diminishing its ability to fight outnumbered and still hold its own. In the extreme case, a 1:1 ratio might be needed. Fortunately, as ratio 5 shows,

[19]Somewhat higher thresholds also are used on occasion (e.g., 3–4:1). The entire subject is clouded with uncertainty; ratios higher than 4:1 are commonly accepted as too high for the defense.

the deployment of Army Group B reserves could help remedy this situation. The contribution of these units would be far greater at higher NATO force levels than at lower ones. With a total NATO posture of 40–48 divisions, the originally adverse force ratio would be pulled downward to 1.25–1.20:1. A posture of 30–36 divisions results in a higher, potentially troublesome ratio of 1.33–1.50:1. At 24 divisions, the ratio remains at 2.0:1 because Army Group B would have no reserves to deploy; the likelihood of defeat in that case would be particularly high.

Army Group A's vulnerability to a breakthrough disaster at low force levels could be partially offset by redeploying northward, for example, four of the 12 Army Group B divisions originally committed to the front line there. This would reduce the Army Group A disparity to 1.20:1 at 24 divisions apiece and would bring the Army Group A balance nearly into parity at higher levels. In an actual crisis, NATO doubtless would attempt this maneuver, regardless of the complexities involved. But consciously planning to do so would take NATO out of the realm of a theater employment doctrine of linear defense as defined here. This analysis therefore will assume that at 24 NATO divisions, no Army Group B reserves would be available.

No matter how many forces NATO decided to extract from Army Group B and redeploy northward to stem the tide in Army Group A, the process would be difficult, time-consuming, and of uncertain effectiveness in the early period. A single division could be deployed within two or three days, depending on its exact location and configuration. Deploying more forces, however, would take longer. A three-division corps would take five or six days; larger forces, longer yet. Also, NATO's support base and supply lines in Army Group A might not be able to surge their output enough to accommodate the demand overload placed on them. The fact that NATO still regards logistics as a national responsibility and has achieved little by way of integrated multinational logistics planning would further complicate matters. A considerable portion of Army Group B's support bases, supply lines, and war reserve stocks would also have to be uprooted and sent northward, imposing additional delays on NATO's ability to get the most out of Army Group B combat forces deploying to Army Group A.

Failure of Army Group B's forces to arrive in a timely fashion could result in their being fed piecemeal into a Soviet meat-grinder and defeated in detail. Consequently, NATO's defense prospects might well hinge on the ability of Army Group A's forces to keep a Soviet breakthrough under control long enough for this Army Group B redeployment to take hold. At all force levels, Army Group A would be outnumbered by 2:1 in forces fighting the breakthrough battle. An important issue is whether operational differences would arise, working for or against NATO, as a function of these force levels. For each force level, Table 32 displays an optimal estimate of NATO's time-phased capability to redeploy Army Group A reserves to the main sector. It also shows how fast Army Group B's reserve combat forces—with combat support and combat service support units but lacking a full reorientation of NATO's logistic system—might be able to redeploy.

This table postulates that NATO's reserves would not all converge on the main sector in short order even assuming a command decision was made on D-Day to begin a theaterwide redeployment. Instead, NATO's reserves would arrive in stages as a function of their circumstances when the order was given and their distance away. NATO's buildup rate in the main sector would rise in a steady, cumulative way rather than all at once. In the best case, reserve units drawn from adjacent corps sectors would arrive within a day. A day later Army Group A reserves would arrive, followed by corps reserves from sectors further away. At the tail end would come divisions drawn out of the line in other Army Group A sectors. Shortly

Table 32

Army Group A's Force Buildup Rate in the Main Sector (additional divisions)

Total posture		48	42	36	30	24
Total Soviet divisions vs. Army Group A		36	30	24	18	12
Army Group A Forces						
Immediately available	(D-Day)	4.5	4.5	3.75	3.0	3.0
Adjacent corps reserves	(D+1/2)	3.0	3.0	1.5	0	0
Army group reserves	(D+2/3)	6.0	3.0	3.0	3.0	0
Distant corps reserves and adjacent front line	(D+3/4)	2.5	2.5	1.75	1.0	1.0
Distant front line	(D+4/5)	2.0	2.0	2.0	2.0	2.0
Army Group B Forces						
	(D+6/7)	3.0	3.0	3.0	3.0	0
	(D+7/10)	3.0	3.0	3.0	0	0
	(D+8/12)	3.0	3.0	0	0	0
	(D+9/14)	3.0	0	0	0	0
Total		30	24	18	12	6

thereafter, Army Group B reserves would begin arriving in a sequential formation dictated by road networks and traffic conditions. This scheme is only illustrative, but it does reflect the basic pattern and flow that probably would guide NATO's response. Army Group A's redeployment would be completed by about D+4 or 5, but depending upon how forces were available, redeployment from Army Group B might take a week or two.

Would this Army Group A buildup rate, at each force level, be adequate to the task of containing a Soviet breakthrough? Table 33 helps provide insights by analyzing Army Group A's buildup rate in each case according to two measures of merit. The first is Army Group A's time-phased ability to bring down the force ratio against all Soviet forces in the main sector to 2:1. The second measure is Army Group A's ability to establish manageable engaged force ratios against Soviet frontline units. It assumes that the Soviets would expand their breakthrough zone enough to funnel an additional three reserve divisions per day into the battle. This equates to a Soviet advance rate of about 15 kilometers per day, a brisk but not improbable rate reflecting what happened at the Battle of the Bulge and well below Soviet doctrinal norms. Both ratios do not reflect attrition. Together, they help shed light on the dynamics of the force balance in the main sector during the period when NATO would be redeploying. A critical issue is whether, during this period, the force ratio might rise too high over a too-long interval for NATO to keep matters under control.

The overall force ratio suggests that four days would pass before NATO could bring the balance close to 2:1. This holds true at all five alternative force levels; during D+2 or 3, the smaller NATO postures seem subjected to somewhat greater stress than the larger postures, but not hugely so. Major differences appear when the ratio of engaged forces is examined. NATO postures of 42–48 divisions would be subjected to far less stress than the smaller postures, mainly because the Soviets themselves would not be able to introduce all their forces into the battle immediately. The process would take time, with the pacing element being the

Table 33

Effectiveness of Army Group A's Buildup Rate (ratios = n:1.0)

Divisions	D-Day	D+1	D+2	D+3	D+4	D+5
			Overall Force Ratio			
48	8.0	5.5	3.1	2.4	2.1	2.0
42	6.7	4.6	3.2	2.5	2.1	2.0
36	6.4	5.0	3.3	2.9	2.1	2.0
30	6.0	6.0	4.0	2.7	2.1	2.0
24	4.0	4.0	4.0	3.3	2.7	2.0
			Engaged Force Ratio			
48	1.3	1.4	1.0	1.0	1.0	1.2
42	1.3	1.4	1.3	1.2	1.3	1.4
36	1.6	1.9	1.7	1.8	1.6	1.8
30	2.0	3.0	2.7	2.3	2.1	2.0
24	2.0	3.0	4.0	3.3	2.3	2.0

rate at which their salient is expanding. At higher NATO force levels, which provide a faster buildup rate in the main sector, NATO would be able to keep the engaged ratio down to a stable average of 1.2:1 during the entire interval. The average, with periodic surges up to 1.4:1, can be a difficult proposition for a defender who has been thrown off his prepared positions, is retrograding, and is fighting meeting engagements against an aggressively advancing enemy. But the imbalance is not so large that the Soviet advance rate is likely to accelerate beyond 15 kilometers per day or that NATO's forces are likely to experience highly adverse attrition dynamics and exchange rates. In this sense, NATO is apparently capable of keeping the battle under control.

As NATO's posture is reduced, however, the situation would steadily worsen in a relative sense. At 36 divisions, the balance constantly hovers around 1.7:1. Between 30 and 24 divisions, it leaps upward to an average of 2.6:1, with periodic surges into the 3–4:1 range. This disparity is a much less manageable proposition. The Soviet advance rate might well accelerate beyond 15 kilometers per day. Equally important, an adverse attrition dynamic and exchange rate process might start taking hold. The risk would be that this process might lead to the cascading mechanics modeled by the Lanchester equations that can result in the outnumbered side being destroyed very quickly, with few losses to the attacker.

Although the engaged ratios are worse at lower force levels, in all five cases an outnumbered NATO would be facing an attacker that, after gaining a early breakthrough, would be maneuvering tactically to prevent NATO from restoring a cohesive line. Success almost certainly would guarantee a rapid NATO defeat and an uncontrollable Soviet breakout before Army Group B reserves could arrive. Conversely, successful restoration by NATO, along the lines of what the allies achieved at the Battle of the Bulge, would help NATO buy precious time and prevent a bad situation from becoming a disaster. Table 34 helps gauge NATO's capacity to reform a cohesive line somewhere to the rear. It dynamically measures the capacity of NATO's buildup rate in the main sector to cover the expanding front on 25 kilometer division frontages. The Soviet advance rate again is assumed to be 15 kilometers per day. For analytical purposes, the assumption is made here that NATO's forces do not conduct a theaterwide retreat, an operationally difficult maneuver that, even if successful, would

Table 34

NATO's Capacity to Cover an Expanding Front
(percent of front covered)

Divisions	D-Day	D+1	D+2	D+3	D+4	D+5
			No Attrition			
48	150	162	217	225	225	200
42	150	162	175	183	183	183
36	125	112	133	125	141	133
30	100	71	83	100	110	100
24	100	71	55	55	68	66
			10 Percent Daily Attrition			
48	150	148	205	215	213	190
42	150	148	169	174	174	160
36	125	105	126	120	135	124
30	100	64	70	88	92	82
24	100	64	44	42	54	50
			15 Percent Daily Attrition			
48	150	144	201	211	208	187
42	150	144	165	161	152	127
36	125	102	114	84	94	90
30	100	54	62	80	82	65
24	100	54	40	38	48	43

undercut the important goal of NATO's strategy of protecting Germany's borders. The tactic of trading space for time is discussed in Sec. 3.[20] Allied forces at the Battle of the Bulge successfully employed exactly this tactic to restore control.

Table 34 provides another useful measure of merit for gauging NATO's capacity to bring a dangerous breakthrough under control. It shows NATO's capacity assuming no attrition, but since loss of forces could adversely affect NATO's coverage capacity, it also shows data for daily NATO attrition rates of 10 percent and 15 percent losses to the major weapons of engaged forces. Both attrition rates are within the range of standard estimates for modern warfare.

A larger NATO posture would stand a considerably better chance of restoring and maintaining a cohesive line than a smaller posture, particularly if high attrition were to take hold either through direct casualties or from the inability of some NATO units to extract themselves from the salient to assume a position along the restored line. At high force levels (42–48 divisions), NATO's buildup rate would be amply large to reconstitute a defense line anytime during the five-day interval, even if high losses were absorbed. At 36 divisions, NATO's buildup is more marginal but still mostly adequate. At substantially lower levels, however, NATO's forces would be spread dangerously thin throughout the entire period. To be sure, they would not be devoid of options: For example, NATO's divisions could extend their frontages by about 50 percent by committing all three of their brigades on the line, but this tactic would leave them without reserves and vulnerable to quick breakthroughs all

[20]A rate of 15 kilometers per day is equivalent to an average of two or three successful tactical atatcks per day; Soviet doctrinal norms call for far higher rates. Western computer analyses tend to show slower rates (e.g., 5 kilometers a day) when the defender has enough strength to maintain a strong cohesive defense. The probability of this kind of defense being maintained increases as a function of the defender's strength.

along the line. And at 24 divisions, even this tactic would fall well short of making up the difference.

The general model used here helps capture important features of the breakthrough process, but like any model that strives for parsimony, it doubtless overlooks several variables that could drive an actual war in a different direction than implied here. Also, the output data generated by this model are sensitive to input assumptions whose value could take on a different form in a war. In essence, things plausibly could go worse for the Soviets and better for NATO than envisioned here (or vice versa). Moreover, even if these data are broadly correct, they guarantee no specific single-point outcome. They do not ensure that NATO would successfully contain a breakthrough at high force levels or that NATO would automatically fail at low force levels.

To the extent that the model and data are accurate, they do suggest that the risks facing NATO would be fairly low at high force levels but would rise as NATO's forces are reduced. With a posture of 42–48 divisions, NATO should be capable of bringing an early breakthrough under control. At least it would be able to build up sufficient forces promptly to restore a cohesive defense line in the main sector and to keep engaged force ratios within manageable limits. NATO also would enjoy a sizable advantage in tactical air support and should be able to maintain a competitive status in the exchange rate process. Its ample force-to-space ratio would enable it to give ground, thereby allowing a larger salient to develop, to reduce attrition. NATO should thereby be able to buy enough time for Army Group B reserves to redeploy northward, achieving a satisfactory 1.2:1 overall balance under manageable conditions. Barring some unexpected reversal, the ultimate outcome would be similar to what occurred at the Battle of the Bulge.

With a NATO posture in the mid-30s, other things equal, the battle would probably be a more closely contested affair, but one that NATO should be able to handle if it uses its forces effectively. Engaged force ratios would remain well on the high side throughout the initial stages, and NATO's tactical air advantage would be less impressive. But NATO should have enough manuever units at least to restore a cohesive line. Its ability to maintain that line long enough for help to arrive would depend heavily on its capacity to keep its casualty rate down while not allowing the salient to expand too far. Initially NATO would have a sufficiently high force-to-space ratio to trade some space for time. The risk would come if as the battle unfolded, NATO failed to manage the interaction between the battlefield's geometry and the attrition process. For example, reluctance to accept a further expansion of the salient could lead NATO to expose itself to a high loss rate that could erode its capacity to maintain a uniformly strong line. The converse would also be true. Short of this kind of failure, Army Group A's forces should be capable of holding out long enough for Army Group B's reserves to redeploy in force. Arrival of Army Group B's reserves would bring the main sector ratio down to 1.33:1, a potentially manageable situation as long as a swirling maneuver battle was not underway.

A less sanguine state of affairs would unfold if NATO suffered an early breakthrough with a posture of 30 divisions or less. Engaged force ratios would be well above manageable levels for any situation short of prepared defense positions, much less a string of meeting engagements with NATO forces on the retreat. Nor would NATO's tactical air forces be capable of lending the kind of support that could be offered at higher force levels. With few, if any, Army Group B forces likely to be available shortly, NATO's prospects might hinge on Army Group A's ability to establish a coherent line, with prepared positions, in the main sector. If the Soviets advanced any appreciable distance into the rear areas, however, the battlefield would have opened up to the point where Army Group A's forces there would lack

the force-to-space ratios to form a contiguous line. NATO might be compelled to fight a pure maneuver war of meeting engagements while heavily outnumbered.

Defeat would not be inevitable in this circumstance, and NATO's forces would have options at their disposal. The conditions under which NATO could succeed would involve some combination of Soviet ineffectiveness, a high NATO proficiency in conducting a mobile defense, and a major NATO advantage in weapons effects and firepower. Other contributing factors could include a reallocation of tactical air strikes away from other missions and a decision to pull sizable Army Group B forces out of the defense line. In World War II, the German Army often salvaged similar situations on the Russian front by fighting a maneuver war more effectively than did their Soviet enemy. NATO could plausibly do the same today.

Short of these factors working together to transform the force-driven dynamics discussed here, NATO's forces would be subjected to major stress. There would be a tangible risk that the Soviets could defeat NATO's forces in detail both in the main sector and then across all of NORTHAG and possibly the entire theater as well. Much would depend upon the complex interplay of battlefield and operational dynamics that are hard to analyze in advance. What can be said is that at low force levels, NATO would have a smaller margin of insurance than at higher levels, and it would face longer odds. Table 35 summarizes these conclusions in capsule form.

Table 35

Key Features of NATO's Defense Prospects at Alternative Force Levels

Force Level	Main Sector Battle in Army Group A	Army Group B Reserves
42–48 divisions	Manageable engaged force ratios Strong air support Reconstitutable line High force-to-space ratio Multiple options for managing battle	9–12 divisions 1.2:1 ratio in main sector
Mid-30s	Adverse engaged force ratios Moderate air support Adequate force-to-space ratios Reconstitutable line, but dependent on effective battlefield management	6 divisions 1.3:1 ratio
24–30 divisions	Unmanageably high engaged force ratios Low air support Inadequate force-to-space ratios Nonreconstitutable line Outnumbered maneuver battle likely	Few, if any, reserves 1.5–2:1 ratio

Interactive Analysis: A Late Breakthrough Scenario

The analysis thus far has considered only a scenario of a shock-induced Soviet breakthrough in the first day or two. What about a scenario involving a later breakthrough, after a period of intense fighting, mutual attrition, and slow movement rates at the front line? In this event, the Soviets, having failed to rupture NATO's line through shock tactics, would have to resort to wearing down NATO's forward defenses in a grueling, time-consuming attrition battle. Once a hard-won breakthrough had been gained, the Soviets would then have to hope that their own surviving forces were still powerful enough to win the rear area struggle against whatever NATO forces had survived.

The nature of the challenge that NATO would face in this situation would be quite different, and probably less stressful, than in the first scenario. There, NATO's task was one of trying to redeploy enough units to reassemble an already ruptured line in the face of a rapidly growing salient. In this scenario, NATO would be endeavoring to prevent a rupture from occurring in the first place. Its frontline forces would have stood up well under the initial shock of combat. By maintaining a cohesive front, they would have denied the Soviets any rapid advance. Although holding their ground initially, they would be subjected to steady pounding from Soviet armor and artillery. The risk is that they would steadily wear down and eventually give way. NATO's task would be one of bringing enough reinforcements to replace losses as they occurred, thereby preventing an already stable line from weakening. Meanwhile, NATO would employ its direct and indirect firepower in an effort to inflict adverse exchange rates on the Soviets, hoping to outlast them in this race to exhaustion.

The time dimension and the exchange rate process would heavily influence the outcome of this scenario. The Soviets presumably would aspire to gain a breakthrough as quickly and inexpensively as possible. NATO meanwhile would aim to extract a heavy toll in a fashion that did not unduly weaken its own forces, and it would have an incentive to slow the pace of battle to buy time to reinforce. The outcome would depend upon which side succeeded in shaping the battle to conform to its own dictates.

In general, battlefield dynamics in this scenario would work in NATO's favor. The Soviets would have the advantage of being able to concentrate along the front and thereby to initiate each engagement on numerically favorable terms. But NATO would benefit from its ability to fight from prepared positions that enhance its survivability and to generate the high firepower rates that can inflict large losses on the enemy. If time was not a factor, the Soviets might be able to surmount the barriers facing them by carefully choosing the times and places to attack, wearing down NATO's defenses at an acceptable cost, but they would not have the luxury of dragging the battle out and therefore would be compelled to accelerate their tempo of operations. This requirement could lead them into a series of hasty, uncoordinated, and ill-focused attacks that might be unsuccessful, too costly, or both.

NATO meanwhile would have greater flexibility and a broader range of options at its disposal. It would probably face a slow but steady Soviet advance into its territory along the main sector. It would have to take care to ensure that the dynamics of this growing salient did not lead to a major rupture of its line. But subject to this constraint, it would have greater scope for picking and choosing its own fights. It could give battle when the circumstances were right and avoid battle when conditions were unfavorable. Lacking any incentive to accelerate the pace of events, it would be under no compulsion to risk valuable forces in the pursuit of uncertain rewards. NATO would be fighting in circumstances that seem conducive to a successfully managed defense effort.

Yet absolute force levels at parity theoretically could enter into the military equation to help push events one way or the other. How exactly would NATO force levels interact with time and attrition dynamics? Table 36 helps address the time dimension by displaying how many days, at each NATO force level, might be required before the Soviets could gain a breakthrough by attrition mechanics alone. It measures how long the Soviets would have to maintain attrition-inflicting pressure on Army Group A's forces in the main sector to wear them down physically to the point where they no longer could populate the front line there on doctrinal frontages. This criterion is four divisions (five divisions would be needed to populate a 75 kilometer main sector with an average enemy advance of 25 kilometers beyond the FLOT).

For each NATO force level, four attrition rates are employed, measuring daily losses—in direct casualties and unit loss of cohesion—to NATO engaged forces and armored weapons in the main sector. Overall theater loss rates would be much lower (e.g., 1–3 percent daily). The estimated day of breakthrough in each case reflects the above criterion. This table is based on a hierarchy of assumptions that are sensitive to numerous input variables. Its results therefore are illustrative, not definitive, and they should be used solely for comparative evaluation. Its input assumptions and data broadly reflect those used in many computer simulations of a European war, and they are employed in a consistently symmetrical fashion.

These data acquire operational meaning when they are compared with the 1–2 weeks that would be required for large Army Group B forces (reserves or frontline) to redeploy to Army Group A. They suggest that NATO would be less likely to face serious time pressures at higher than at lower force levels. A NATO posture of 42–48 divisions would have one or two months of staying power at the lower attrition rates, and 2–3 weeks at the higher rates. A NATO posture in the mid-30s would have 3–4 weeks of staying power if casualty rates were low, but a more marginal 1–2 weeks if they were high. At 30 divisions or less, Army Group A's forces plausibly might be worn down and broken through before help could arrive. This would be the case particularly if the Soviets succeeded in accelerating the pace of attrition combat.

If the Soviets did gain a breakthrough by attrition mechanics, they would need a sizable force to exploit the opportunity in NATO's rear areas. Postulate that the Soviets would require an overall 1.25:1 advantage or more against remaining Army Group A reserves plus Army Group B reserves that would be either at the scene or converging on it. Table 37

Table 36

Days of Combat Before Soviet Breakthrough Can Be Achieved (days after D-Day)

NATO Divisions	Daily Attrition Rates (percent)			
	5	10	15	20
48	65	32	21	16
42	50	25	16	12
36	35	18	11	9
30	20	10	6	5
24	5	3	2	1

Table 37

Postbreakthrough Force Ratios in Main Sector

	Soviet/NATO Exchange Ratios (n:1.0)		
NATO Divisions	1:1	1.5:1	2.0:1
48	1.35	.98	.59
42	1.43	1.07	.71
36	1.54	1.27	.91
30	1.75	1.50	1.25
24	2.20	2.10	2.00

displays post-breakthrough force ratios as a function of alternative NATO force levels and exchange ratios in the main sector. The spectrum covered is from 1:1 to 2:1 favoring NATO. A reasonable best estimate is 1.5:1 in NATO's favor.

The table suggests that force levels would play a role in determining NATO's prospects. At higher force levels (42–48 divisions), the Soviets would be unable to exceed the threshold requirement for prosecuting a breakthrough unless they were successful in achieving equal exchange ratios as they were fighting their way through NATO's prepared defenses. With a NATO posture in the mid-30s, the Soviets would meet (although not greatly exceed) the threshold if the exchange process unfolds according to the best estimate. NATO could prevent the ratio from rising this high only if it were to achieve favorable exchange rates of 2:1 or better. At 30 divisions or less, NATO's prospects would worsen: Even a 2:1 exchange ratio might not suffice. At 24 divisions, for example, a favorable exchange rate of 3.5:1 would be necessary for NATO to bring down the postbreakthrough ratio to 1.25:1.

Taken together, these two tables permit the following observations about how NATO force levels would interact with time pressures and attrition dynamics in this scenario. At high force levels, NATO's posture would be a very tough nut to crack. To gain any serious prospect for breaking through, the Soviets would have to drive up NATO's daily casualty rate to 20 percent or more while simultaneously holding down the exchange rate to about 1:1. This task would be exceedingly difficult against an Army Group A posture that had both prepared positions at the onset plus so many reserves that it could give ground locally to keep its own losses under control. In all likelihood, the Soviets would suffer adverse exchange rates of 1.5–2:1 or worse. In that event, the Soviets probably would be capable of continuing their attack only a few days, after which they would have to break off and swing over to the defensive to stem their losses before NATO gained an exploitable advantage of its own. A stalemate after about a week of fighting, long before NATO's posture had been worn down to the point of breaking, would be the most likely outcome.

A war fought at 36 divisions apiece would be a closer proposition, but one that NATO would have the immediate resources to manage. Since NATO's posture would have less depth and staying power, the Soviets would face fewer pressures to accelerate the pace of combat in ways that could result in unacceptably high losses to themselves. In addition to being able to attack in a more controlled fashion, they would face a less demanding exchange rate calculus. The data here suggest that the critical issue would be whether they could drive up NATO's daily loss rate to about 15 percent or somewhat more, while keeping the

adverse exchange rate down to 1.5:1 or less. This itself would be no easy task, but both parameters do fall within the range of plausibly achievable estimates.

The quality of NATO's defense effort would be the telling factor. By achieving favorable exchange rates of 1.8:1 or better, Army Group A's posture could hope to stalemate the attack itself within about a week. But if Army Group A's performance dipped well below this standard, the battle could be sharply contested. There would be a continuing risk that Army Group A, in attempting to balance attrition and the FLOT movement rate, might stubbornly hold onto positions at an unwisely high cost. Alternatively, it might retreat too early and too often, thereby causing the front line to expand beyond NATO's capability to cover it in adequate density. But Army Group A would have enough reserves to provide flexibility and a range of tolerance for some tactical errors. Provided the defense effort was reasonably competent, Army Group A's forces should be capable of holding on long enough for Army Group B's reserves to arrive in strength.

NATO's defense prospects at 30 divisions or less would be more problematic in the absence of reinforcement from Army Group B frontline forces. Army Group A's force-to-space ratio would be so low that the Soviets realistically could aspire to gain a major breakthrough within a few days. Also, they could attack in the confidence that even if they absorbed moderately adverse exchange rates in this period, they would still have available a surviving exploitation force. Army Group A's forces would need to achieve favorable exchange rates of 2.0–3.5:1 to stalemate this attack. These are not implausible performance standards, but they could be hard to achieve with a posture that had little flexibility for giving ground locally because the frontage demands of an expanding salient could not readily be met. At this low force level, the time dimensions of an attrition-driven breakthrough converge with those of a shock-induced scenario: In both cases, breakthroughs could come within a few days. NATO's prospects would again depend heavily on the ability of Army Group A's forces to combine fire and maneuver in a way that compensates for a lack of mass.

Having said this, it also is true that the outcome in this complex scenario would be subject to many variables whose interplay is more complicated than any verbal assessment can hope to convey. This scenario was subjected to analysis with both the IDAHEX and CAMPAIGN dynamic simulation systems. The results from both models called attention to the major uncertainty confronting an appraisal of this topic and to the sensitivity of all outcomes to changes in assumptions and data. Within this framework, both simulations tended to confirm the basic proposition that at some point along the force reduction continuum, a successful NATO linear defense might well give way to failure in this scenario.

The IDAHEX system was employed to conduct rather detailed war gaming analyses of NATO's performance with between 40 and 30 divisions (DEFs). Employing assumptions that broadly reflect the mainstream of many Western simulations, the results showed a consistent stalemate battle at 40 divisions apiece. As mutual reductions were taken below this level, the stresses on NATO's posture steadily mounted. NATO's task started becoming unmanageable at about 35 divisions. At 30 divisions, the Soviets regularly were capable of gaining an exploitable breakthrough. The CAMPAIGN simulation system produced results showing NATO's linear defense holding at levels above 33 DEFs (27 EDs). Once NATO's posture was reduced to the range of 22–26 DEFs (18–21 EDs), the linear defense was regularly penetrated.[21]

There are many uncertainties, and these two simulation systems were designed for different purposes; what stands out is the similarity between the conclusions that were reached.

[21]Davis et al., 1989.

The somewhat different wars modeled by these two simulation systems are both within the range of plausibility and common acceptability. NATO's prospects at any given low force level (below 35 divisions) are not fixed in some deterministic way. As a result, there is no single point above which success is guaranteed and below which failure is inevitable. NATO's prospects would depend heavily on exactly how battlefield dynamics might flow. The more favorable the conditions, the lower the NATO force level at which linear defense would still be feasible. Conversely, unfavorable conditions that work against NATO would tend to elevate this threshold to higher force levels.

The results of these two computer simulation systems help bound the range of uncertainty regarding the viability of linear positional defense, but the range that they postulate is not chaotic. Rather, it is a product of analytically estimable variations in the factors that seem likely to drive the combat process. Also, this range is fairly narrow. Other things being equal, somewhere within this range deep cuts apparently would make NATO's linear defense doctrine quite difficult to execute, perhaps unmanageably so.

SENSITIVITY ANALYSIS

The viability of NATO's linear defense doctrine would also be affected by a host of combat-influencing factors that lie outside the boundaries of the general model developed here. These outside factors can be called "exogenous" variables, as opposed to the model's "endogenous" variables. Some of them would work in NATO's favor, others against. Perhaps the most important variable is parity itself. In essence, a minor Soviet numerical advantage would matter more against a small NATO posture than against a larger one. Four other variables will be subjected to a sensitivity analysis:

- Differences in relative reinforcement rates.
- Logistic constraints on each side's performance.
- Command and control dynamics in the era of modern technology.
- The added firepower provided by future deep strike FOFA systems, including tactical air forces and ground-based missiles.

Relative Reinforcement Rates

The analysis thus far has assumed that at each force level, both sides would have been able to fully deploy all their forces to the battlefield by D-Day. This assumption, however, oversimplies a more complex reality. Most likely a Central European crisis would begin with a competitive mobilization and reinforcement race between the two sides. Fighting probably would break out before one side or the other had completed its buildup. Differences in relative buildup rates therefore could have an important bearing on what otherwise would be a parity battle by temporarily giving one side or the other a numerical edge.

In general, relative buildup dynamics would work in favor of Soviet forces by virtue of the USSR's close proximity to Central Europe and the far more distant basing of many U.S. forces across the Atlantic Ocean. Degree, however, matters greatly here. In past years, NATO's defense plans have been based on the assumption that the Soviets could fully mobilize a large force and deploy it to the forward areas within 15–30 days. This left NATO capable of drawing only on those U.S. reinforcements whose equipment already had been prepositioned in Europe, some six additional divisions. Political realities in Europe now are changing in ways that

would sharply constrain the Soviet Union's capacity to quickly deploy large forces to the forward areas. As long as sizable Soviet forces remain on Germany's soil, a reinforced Soviet attack could conceivably be launched as early as M+30. But once Soviet forces are fully withdrawn from Germany, scenarios involving a Soviet attack any earlier than M+60–90 seem unlikely. This elongated Soviet buildup rate should take some of the pressure off NATO's own mobilization effort and reduce the dangers inherent in relying on outside reinforcement from the United States.

With more time to mobilize, NATO should be better able to complete the process of sending not only the prepositioned U.S. units to Europe but other active forces as well. Nonetheless, NATO's posture will remain dependent on outside U.S. reinforcement, and more so if NATO were to retain a linear employment doctrine after CFE I is implemented and other force cuts have taken effect. Table 38 helps illustrate the effects by displaying four NATO buildup rates: the current capacity, after CFE I's 10 percent cuts have taken effect, after a hypothetical 25 percent cut, and after a 50 percent drawdown.

The table suggests that after CFE I drawdown, which would leave about 40 divisions at M+30, NATO would have sufficient forces throughout the mobilization process to contemplate a linear array at parity. Moreover, NATO would not be highly dependent on a rapid U.S. reinforcement effort to establish a reasonably confident linear array. Force cuts of about 25 percent would begin to place strains on NATO's posture. The Alliance would have only about 30 DEFs in place in Central Europe. It thus would be left dependent upon the U.S. reinforcement effort; deployment of the prepositioned U.S. III Corps, which provides AFCENT's principal theater reserves, would be particularly important. Force cuts of 50 percent would leave NATO with only some 20 DEFs in Central Europe, well below what all of the methodologies employed here suggest is needed for a linear defense even at parity. Even after full U.S. reinforcement, moreover, NATO would be capable of deploying only 29 DEFs

Table 38

Effect of Force Reductions on NATO's Buildup Rate (DEFs)

Reduction	M+30	M+60	M+75
Programmed (1989)			
In place	40	40	40
U.S. reinforcement	5	14	17
Total	45	54	57
Post-CFE I			
In place	35	35	35
U.S. reinforcement	4	14	17
Total	39	49	52
After 25 percent cuts			
In place	30	30	30
U.S. reinforcement	4	11	13
Total	34	41	43
After 50 percent cuts			
In place	20	20	20
U.S. reinforcement	3	7	9
Total	23	27	29

in total. This posture would be on the short side of adequacy by most of the methodologies used here.

Logistic Constraints

All assessments of Central Region ground force combat are sensitive to the ability of both the attacker and the defender to provide adequate levels of combat service support to their maneuver formations. This has been the case in the past and will remain true in the future. How then would logistic support factors affect NATO's linear defense prospects at parity? Although only a speculative appraisal can be offered here, logistic factors would apparently affect both sides in ways that work both for and against NATO. The balance between these competing dynamics is difficult to assess with any precision, but they do create strong incentives for NATO to pursue multinational integration of its logistic support forces if its defense plans are to employ the counter-concentration tactics that appear necessary for linear defense to work at lower force levels than now.

Working in NATO's favor is the fact that the Soviets historically have provided considerably smaller combat service support forces for each of their ground divisions than has NATO. As a result, NATO's larger logistic support forces—relative to the number of divisions being supported—often have been regarded as one factor bringing the Central Region balance into closer alignment than surface appearances have suggested. Other things being equal, parity in ground combat forces would leave NATO with a sizable superiority in service support forces of around 2:1 in manpower. To the extent that these extra assets can be translated into actual combat power, this edge appears to give NATO a hidden margin of insurance at parity. In principle, it also might help buffer against the risks of conducting a linear defense at low force levels.[22]

Weighing against a whole-hearted acceptance of this conclusion are two countervailing considerations. First, it always has been unclear exactly how the Soviets' leaner support structures would have impeded their capability to prosecute their offensive military doctrine. The Soviets have long sought to fight and win a short war in Central Europe, one in which combat forces would predominate and support forces would not play a dominant role. In theory, logistic support forces can matter heavily even in quick, violently prosecuted breakthrough operations. Although Western military authorities have recognized this reality, the Soviets themselves obviously have not taken their support shortcomings seriously enough to alter their force structures. In the years ahead, it is possible that as their combat forces are reduced, the Soviets will elect to beef up their logistic support units. There are no constraints stemming from East-West arms control accords on their freedom to take this step. Regardless, it is unsafe to plan NATO's defenses on the assumption that the Soviets would leave themselves weak in this important area.

It is equally unclear how any NATO advantage in logistic support would translate into a militarily decisive edge in the kind of war that would be fought at low force levels. Although a long war cannot be ruled out, this war probably would be decided in the initial stage of operational maneuver, and within the first week or two. Many of NATO's logistic support forces would play a critical role in the outcome, but others would provide sustaining support and staying power that would be manifested only over a longer period of time.

Even those support forces that did figure heavily in the defense equation could not perform the military missions of combat forces. They could help each of NATO's ground

[22]Department of Defense, 1984.

divisions fight harder and more effectively by enabling them to move faster, expend ammunition at higher rates, and fire more accurately. But they could not plug holes in the front line, block enemy penetrations, or conduct counterattacks. They primarily would enhance the quality of NATO's posture, not its quantity. The problem with taking major credit for any NATO superiority in logistic support is that NATO's linear defense could get into trouble at low force levels precisely because of its lack of quantity.

Logistic support dynamics could contribute to NATO's undoing. Although the Soviets have not fielded large support assets, they have benefitted from having a uniformly constituted and fully integrated force posture that has made logistic support a fairly easy task for them. This advantage doubtless has enabled them to avoid duplication and redundancy, thereby allowing them to get by on less.

Operating under a philosophy that declared logistic support to be a national responsibility, NATO has been compelled to provide separate support structures in Central Europe for seven national armies. Lacking the capacity to draw on NATO-wide assets, the support structures for each army had to be designed to function in high intensity combat even though NATO's corps were unlikely to have come under heavy pressure in an actual war. With each army having its own unique weapons and doctrines, a substantially larger support structure was required than might otherwise have been the case. Along with efforts to meet these requirements have come redundancy and duplication that thus far have stubbornly resisted efforts to correct them. To some degree it is these factors, rather than any optimal theory of combat power, that accounts for NATO's large support forces.

As a result, NATO's layer cake today is backed up by what can best be called a "piston" logistic system. That is, each forward corps sector is backed up by a national support structure that is quite different in composition from its neighbor on either side. Some of these logistic pistons have a surge capability built into them, enabling them to support additional combat forces for a time. But none is designed to supply forces from nations other than their own. This piston structure retards NATO's flexibility in shifting its combat forces from one area to another along the front, compounding the problems of trying to conduct a linear defense at low force levels.

Earlier, this study discussed how this inflexibility might constrain NATO's capacity to shift forces from Army Group B to Army Group A. This piston logistic system might also impinge on NATO's capacity to redeploy forces within Army Group A itself. The scenarios discussed here, for example, envision the counter-concentration of 6–18 Army Group A divisions within the immediate area of a 75–100 main sector there. Given Army Group A's multinational coloration, no more than about four of these divisions could come from any single national army. To assemble a force of 12 divisions, units from at least three and possibly four armies would have to converge on the scene.

Whether their somewhat different doctrines and weapons would enable these forces to conduct a coordinated defense is itself a troublesome imponderable, but logistic support problems seem almost destined to compound their difficulties further. Quite apart from the matter of overload, the purely qualitative capacity of local logistic units to support them would be questionable. As a result, NATO's rear support forces would have to reorient and redeploy in a manner similar to their Army Group B counterparts. At a minimum, the entire process might delay the redeployment of badly needed combat forces and otherwise take more time than NATO had at its disposal. Even when this reorientation is completed, NATO's supply lines would resemble a complicated patchwork with multiple seams overlapping each other. These very real problems create compelling reasons to achieve greater

multinational integration of NATO's support posture. Without this kind of solution, they hardly augur well for a linear defense that might lack immediately available combat forces.

Command and Control

One of the hallmarks of traditional defense analysis has been to grant the attacker the advantage of concentrating before the defender can begin redeploying his own reserves to the main sector. The primary reason for this practice is that the defender traditionally has lacked enough intelligence information on enemy deployments to understand the battlefield before the attack begins. The ongoing deployment of highly sophisticated all-source intelligence systems (e.g., JSTARS) is helping to reduce this liability for the defender by giving him greater advanced warning of how enemy concentrations are shaping up at any one time. Once fully in the hands of NATO's commanders, these systems, in principle, should enable NATO to begin redeploying even before an attack begins. If fully borne out, this important advantage would help ease the task of conducting a linear defense at low force levels by enabling NATO to get a quicker jump on its redeployment problem.[23]

To some extent, this analysis already has accommodated these improvements by assuming that NATO would have sufficient clarity to begin a wholesale, theaterwide force redeployment only hours after an attack is launched. This assumption accounts for a critical feature in the calculations presented earlier: that all NATO reserves would begin moving quickly and that they would all converge on the right spot in a period elongated by nothing more than the physics of moving. These seem like fair calculations in the modern era. The question is whether NATO's redeployment efforts will be improved even further by modern collection systems.

Several constraints could stand in the way. The Soviets might succeed in blinding the eyes of NATO's collection systems. For example, JSTARS aircraft would fly close enough to the battlefield to place them within range of Soviet interceptors and possibly long-range SAMs. Other NATO systems have potential vulnerabilities of their own: Satellites can be shot down, radars jammed, and electronic receptors sent confusing messages. Intelligence collection in future wars might continue to be a less than perfect enterprise. Even if perfect technical information is available on the enemy's order of battle, the problem remains of processing this information, reading it accurately, and then drawing the correct conclusions. NATO has a potential problem in this area because the task involved is more than intellectual; it is also organizational. Unfortunately, NATO's intelligence command and control system is itself complicated and not fully integrated. This liability can produce delays, conflicting assessments, and even wrong judgments as intelligence data pass up the chain of command and are synthesized into an overall appraisal.

Another potential constraint is that the Central Region battlefield might not be easy to read even if perfect intelligence is both available and interpreted correctly. The problem would be less one of discerning where enemy forces are located at any particular time than forecasting where they are likely to be deployed in the future, perhaps several days hence. The task of making these forecasts in Central Europe might be a complicated one, especially if the Soviets chose to employ feints and other confusion-inducing tactics. The presence of a reserve Soviet tank army directly opposite one NATO sector, for example, would be no guarantee that this army might not strike against a neighboring sector. Moreover, enemy reserves would be capable of moving quickly: The Soviets would match NATO's capacity to move divisions 50–100 kilometers per day across open ground.

[23]See Pierre et al., 1986.

NATO command decisions on redeploying forces therefore would be more complicated than simply reading the enemy's geographical disposition and realigning NATO's posture on the assumptions that this disposition will remain fixed. Relative motion dynamics would need to be taken into account constantly as both sides struggled to shift their forces to gain the upper hand. These dynamics do not imply that NATO's decisions inevitably would be wrong or that modern intelligence systems could not ease the task enormously, but even with these systems, the risk of indecisive paralysis or erroneous moves that could not quickly be corrected would always be present.

The Soviets themselves already have deployed some of their own modern collection systems and will be acquiring others in the future. To some degree, this development will checkmate NATO's advantage in this area. Although the Soviets will no longer be capable of concentrating secretly against one segment of NATO's line, NATO itself will no longer be capable of either disguising its prepared positions as easily as before or carrying out its own counterdeployments in secrecy. In essence, both sides will be able to conduct their moves with greater confidence that they are properly focused but with less assurance that the other is likely to be caught off guard. It is difficult to tell whether this revolutionary development will work to the benefit of the attacker or the defender.

FOFA and Lethal Operational Fires

Another revolutionary development is the impending appearance of highly lethal systems capable of striking deep (e.g., 100 km) against pinpoint targets, including armored vehicles. Included in this category are the U.S. Army's ATACMS missile mounted on the Multiple Launch Rocket System platform, with the missile dispensing antiarmor bomblets guided by modern sensing devices. Also included are equivalent Air Force munitions, mounted aboard aircraft. Rounding out this entire FOFA system is the JSTARS collection platform, which is to provide real-time intelligence information on enemy targets. The idea behind FOFA is to employ these assets to disrupt and degrade enemy reserve armored forces as they approach the battlefield. Alternatively, FOFA systems can be fired across neighboring NATO corps sectors, thereby providing long-range indirect fires.

In principle, FOFA's long-awaited appearance on the battlefield promises to ease the problems of defending linearly at low force levels by enhancing the firepower of NATO's remaining ground and air units. This added firepower would have two positive effects. First, it might enable NATO to impose additional movement delays on echeloned Soviet reserve forces approaching the front, thereby giving NATO more time to constitute its defenses. Second, it should enable NATO's forces to achieve more favorable exchange ratios, reducing the pressure that attacking Soviet forces could place on them.[24]

A simple example will indicate how FOFA might affect the firepower balance. Assume that an existing NATO posture of 30 ground divisions and 2200 aircraft has the combined capability to destroy some 600 enemy armored fighting vehicles per day on D-Day. This equates to 2 percent of the enemy's armored posture. Further assume that NATO could expect to lose 400 armored vehicles of its own each day, thus producing a daily 1.5:1 exchange ratio in NATO's favor. In principle the acquisition of FOFA air and ground systems might be capable of increasing NATO's lethality by, for example, as much as 50–75 percent. This would shift the exchange rate to a more advantageous 2.25–2.6:1. Equally important, it would reduce the Soviet capacity to maintain steady offensive pressure over a sustained period. For example, if

[24]Pierre et al., 1986.

their previous capacity was to sustain an offensive for 15 days in the face of NATO's firepower, they would now be compelled to halt in only 8–10 days because their losses had passed maximum limits.

Firepower enhancements of this magnitude clearly could transform the battlefield in NATO's favor at parity and make linear defense a still viable option at lower thresholds than otherwise would be the case. At the same time, real-life constraints are likely to arise that could deny NATO the full range of these benefits. One consideration is that FOFA systems are not yet fully deployed and might fall victim to the budgetary axe in the fiscally constrained times ahead. Another concern is that technologically sophisticated systems like FOFA, that have not yet been tested on the battlefield have in the past broken down when used for the first time. A third possibility is that FOFA requires a highly skilled military organization to employ it effectively; the lack of previous combat experience could result in performance degradations. Finally, the Soviets will probably develop suppressive techniques of their own that could blunt FOFA's sharp edge.

Even if these constraints do not come to pass, an offsetting development looms ahead: Eventually the Soviets themselves will acquire a FOFA system of their own. Indeed, the act of cutting back the size of the Soviet Army might free funds for research, development, and modernization, allowing the Soviets to close this and other qualitative gaps quicker. The appearance of a Soviet FOFA would have an uncertain effect on NATO's linear defense doctrine. Both sides would be able to project their ground firepower across a broader distance and deliver more lethal air support within current range-payload limits. This could grant any NATO ground posture a capacity to defend a larger segment of terrain by means of firepower, although not by mass. This change could make linear defense easier.

Mutual deployment of FOFA could, however, accelerate the pace of attrition on the battlefield while holding the exchange ratio constant. A faster battle might work in favor of a concentrated attacker by accelerating his capacity to destroy defenders in his way. This presumably would enhance his ability to gain a breakthrough before the defender had time to redeploy his reserves. Once a breakthrough had been achieved, the defender would face constraints on employing his operational fire systems to close the breach because the attacking and defending forces might be too intermingled to tell apart from afar. Whatever the case, FOFA's appearance clearly would enhance the attacker's incentive to use shock tactics to unravel the defender's front wall quickly, avoiding a prolonged and costly attrition contest. To the extent that better shock tactics or FOFA itself enhanced the attacker's ability to penetrate more quickly, the presence of operational fires might make linear defense even more difficult than before.

SUMMARY

The viability of NATO's linear defense doctrine is likely to be affected by many factors, including but not limited to NATO's own future force levels. The task of trying to assess how these multiple factors will interact is a difficult undertaking, one that can only end with the acknowledgment of many uncertainties. Nonetheless, it is possible to cut through the fog of confusion to offer some policy-relevant observations regarding how force reductions are likely to affect NATO's theater employment doctrine.

The central conclusion of this analysis is that the risk of a linear defense being shattered by an offensive strategy of breakthrough and maneuver is worrisome enough to be taken seriously in NATO's defense planning. History alone supports this judgment, as does a

careful analysis of battlefield motion dynamics employing prudently conservative assumptions. Moreover, force levels do play a contributing role in this equation. A linear defense at low force levels would depend upon the capacity to counter-concentrate rapidly, but its very lack of reserve forces would impede this capacity. In other words, it would lack the medicine to heal itself.

It is impossible to identify any single threshold below which linear defense is impossible and above which success is guaranteed. Just as clearly, larger NATO forces buy greater insurance and confidence. Whether these returns are sufficiently compelling to justify the additional cost of larger forces is a matter that can be decided only in the framework of a larger political and budgetary calculus than employed here. Even at parity, the security returns that would accrue to NATO between postures of 25 and 40 divisions are more than marginal in the absence of a change in NATO's theater employment doctrine.

Conceivably, the development of unilateral NATO advantages in military technology, coupled with a remedying of NATO's logistics problems, could prolong linear defense's life. Barring these developments, the prospect of sizable cuts in NATO's ground combat forces does seem likely to increase the stresses that NATO would experience in trying to implement a linear defense on a lengthy front against a skilled Soviet attack. This particularly will be the case if, as seems likely, the future force balance produces not parity but numerical superiority for the Soviets. A NATO linear defense with 40–45 divisions may be viable if the enemy is similarly sized, but not if the balance is 1.25–1.50:1 against NATO. In this event, linear defense has little choice but to give way to the kind of echeloned defense that has more reserves and can react faster. This need for a new employment doctrine will remain if NATO's forces are reduced further, even if arms control negotiations simultaneously bring the force balance into parity.

Table 39 illustrates how an effective employment doctrine can help improve NATO's defense prospects in situations of either parity or a marginal imbalance. Under a linear defense doctrine, NATO's prospects are mixed at parity and unfavorable if NATO is marginally outnumbered, with NATO's density affecting the outcome in both cases. With an echeloned defense, NATO's prospects improve; in particular, NATO has a reasonable chance for success even if it is somewhat outnumbered.

The alternative to doctrinal reform is unpleasant prospects ahead. Although the Cold War has been officially proclaimed over, a still-vulnerable NATO conventional posture would

Table 39

NATO's Defense Prospects as a Function of Employment Doctrine

	Linear Defense		Echeloned Defense	
Free Balance: NATO's Density	Parity	Marginal Imbalance	Parity	Marginal Imbalance
Higher	Success	Probable Failure	Success	Success
Medium	Uncertain	Failure	Success	Probable Success
Lower	Probable Failure	Failure	Success	Uncertain

not be inconsequential even in the presumably more harmonious post–Cold War era. In essence, the Alliance's successful efforts to achieve parity would not have brought genuine military stability in their wake. To the extent that the Alliance's defense posture remains vulnerable, NATO's deterrence and defense goals will be commensurately weakened. This state of affairs might not matter if relations with the Soviet Union continue improving, but it could become important if these relations were to unexpectedly turn sour. In addition to remaining militarily vulnerable to a conventional defeat, NATO would be left dependent upon nuclear weapons to provide the extra margin of necessary deterrence. By consensus, NATO has agreed to modify its strategy to relegate nuclear weapons to a position of last resort. A vulnerable conventional posture would weaken, if not undercut, the fundamental principle of this important strategy change. Beyond this, NATO's internal unity, which still is an important consideration, would hardly be well-served. In particular, Germany, only recently unified amidst a contentious debate about whether it should remain in NATO's fold, would find its strategic position rendered less secure than otherwise would be the case. This development would not augur well for NATO's ability to continue providing the security umbrella that makes continued membership a desirable commodity for that nation.

6. CONCLUSIONS

Time will tell whether the theater employment doctrines developed in Sec. 3, or ones like them, are to be incorporated into NATO's conventional defense strategy. Much will depend upon how Europe evolves politically. This study has assumed that Europe is headed toward an entirely new security system, one in which defense issues, albeit in more muted form than during the Cold War, will still confront the West. The emergence of an even more stable, tranquil European order, or an impotent Soviet Army, could allow NATO to relax its defense guard beyond what is envisioned here. Conversely, any downward slide into post–Cold War instability could propel defense preparedness further into the forefront.

Barring the emergence of a fully cooperative system, the political-military issues surrounding NATO's future theater employment doctrine seem destined to be on the West's agenda in the years ahead. Events have overtaken the old doctrine of linear defense at the inter-German border (or elsewhere) and relegated it to the ashcan of Cold War history. This will be the case even if NATO's arms control successes go well beyond the CFE I Treaty. History, formal analysis, and common sense support this conclusion. If NATO is to adapt to the conditions ahead, it has no alternative but to confront choice and change.

Until NATO designs a new employment doctrine, it will be unable to fulfill the London Summit's call for a new post–Cold War military strategy. Nor will it be able to translate into practical terms the London Summit's call for a less-primed but still-capable force posture that provides insurance and flexibility. Such important issues as NATO's mobilization rate, training and exercise plan, the active/reserve mix, multinational formations, and future roles and missions cannot be settled in the absence of an employment concept. Beyond this, NATO cannot relegate nuclear weapons to a position of last resort until it is first confident that its conventional defenses are in order.

Important budgetary decisions also ride on NATO's ability to fashion a new doctrine. The Cold War's end will result in a downturn of NATO's defense efforts, a trend that the London Summit has officially blessed. But degree matters here. Until NATO defines its new doctrine, it will be unable to specify how many forces are required for the coming era. As a result, it will be unable to determine how far it safely can afford to reduce forces and cut budgets. It will also be hard-pressed to establish program priorities in a coherent way. NATO therefore will run the risk that its budgetary decisions will be made in the absence of vision to guide them.

The need to design a new doctrine stems partly from NATO's quest for a stable military balance of power and security architecture in Europe. The passing of the Cold War does not mean that history has come to an end there. It may merely mean that Europe is repeating a familiar cycle as one era of security affairs gives way to another. Perhaps the Europe of the 1990s and beyond will be more democratic, stable, and integrated than in the past. But barring the Soviet Union's transformation into a benign democracy, the interstate rivalries and instabilities ahead call for NATO to maintain a military counterweight.

The logic of superpower competition is not the only factor in the equation. NATO also will need a new employment doctrine to preserve its own unity. In the excitement of the Cold War's demise, it is easy to lose sight of the role that military strategy plays in preserving NATO's cohesion. And it is equally easy to forget the role that NATO's cohesion plays in promoting Europe's stability.

This failure to adopt a new, eastward-looking employment doctrine could leave Germany uncomfortably exposed to the risks of the post–Cold War era. This unhealthy state of affairs would come at a time when factions within that nation are increasingly questioning Germany's future role in NATO. The consequences—a German policy of self reliance—is as predictable as it is troublesome. Neither NATO's members, nor Europe as a whole, nor Germany itself can afford to risk driving Germany out of the Alliance by NATO's own failure to meet its security obligations.

Similarly affected could be the allegiance of other NATO members. These nations face less compelling incentives than Germany does for extending NATO's doctrine eastward, but they do have powerful reasons for wanting a doctrine that makes military sense, regardless of where it is executed. A vulnerable doctrine, especially one accompanied by a disgruntled Germany that provides one-third of NATO's forces, would be an unattractive commodity. It could only produce mounting doubts about NATO's raison d'être. This applies not only to Western Europe, but also to the United States, whose willingness to provide extended nuclear deterrence is predicated on NATO's ability to maintain a viable conventional defense option.

Just as designing MC 14/3 and the layer cake was difficult, the challenge ahead will not be easily mastered. More is involved than military operations. Learning how to mobilize and fight differently will also require uprooting changes in NATO's internal political arrangements, program priorities, and budgetary policies. It will affect the fundamentals of how NATO functions as a coalition. Fortunately, the task is manageable, if NATO harnesses its intellectual resources and makes the proper politial decisions.

Systematic analysis can help address the planning problems ahead. It is no cure-all, nor is it a substitute for judgment, but it can help to illuminate the alternatives, their complex tradeoffs, and their consequences. The challenges facing analysis, however, will be a great deal different in the post–Cold War era than in the past. The discipline will need to respond or lose relevance.

Just as NATO's defense plans will be uprooted, modern military theory is destined to undergo an upheaval of its own. During the Cold War, Western military theory in Europe concentrated on managing a linear defense doctrine in a bipolar security system. The task of the post–Cold War era will be developing a nonlinear defense doctrine, with a greater emphasis on mobile warfare, in a more multipolar setting. In some ways, military theory in Europe will find itself facing problems reminiscent of those of the late 19th century, when armies were mobile, crises complex, and coalitions fluid.

Operations research will still play a role in analyzing these problems and thereby contributing to military theory, as well as to NATO's policy decisions, but it will no longer reign supreme in the ways that characterized the Cold War's waning years. Analysts will need to employ conceptually broader methodologies. Systems analysis, with its ability to consider multiple objectives, will need to be resurrected. Equally important, analysts will need to develop a more refined capacity to relate military issues to the larger political context. Political affairs and military doctrine will no longer be divorced in the way that prevailed during the Cold War's last years.

Whatever role formal analysis comes to play, sound planning of a military and political nature is the key to NATO's ability to navigate the uncertain, possibly troubled waters ahead. Planning played a big role in helping the West emerge from the Cold War in far better shape than seemed possible 40 years ago. It is an equally valuable instrument for coping with the post–Cold War era. As Eisenhower said, plans are nothing, but planning is everything.

POSTSCRIPT

Research and analysis for this study was completed in February 1991. At that time the Soviet Union had agreed to withdraw its military forces from Eastern Europe, but communism remained in power in that nation, and the Gorbachev regime seemed likely to retain a large army. The shattering events of late 1991 not only swept away communist rule but also destroyed the Soviet state, leaving in its wake a Russian nation and a host of now-autonomous republics grouped together under the "Commonwealth of Independent States." These developments obviously uproot some of the assumptions that were used in this study, but the underlying issues raised about NATO's employment doctrine in Central Europe will remain on the alliance's agenda. Indeed, the defense planning task facing the alliance is likely to be even more complex than it was in 1991.

For the future, a core issue will be whether NATO should plan its defenses against the contingency of Russian/Commonwealth military aggression in Central/Eastern Europe. At their summit in late 1991, NATO's leaders made clear their desire to build a cooperative relationship with Russia and its Commonwealth partners, but they also expressed the belief that NATO must preserve a military balance of power with its former adversaries in Europe. This step suggested that NATO's defense policy would continue to be anchored on conservative principles, allowing room for this contingency as one element of the force planning process. Further evidence came early in 1991, when press reports indicated that the U.S. Department of Defense was employing, as part of its planning, a scenario involving warfare with Russian/Belarus forces in Central/Eastern Europe.

From what can be ascertained today, the fragmentation of power taking place in the former Soviet Union makes it unlikely that Russia will be capable of generating as many as 60 divisions for any power projection mission beyond its borders. A more plausible projection force would be composed of 25 to 30 divisions and about 1000 combat aircraft. If this force were properly trained and equipped, however, it could possess important offensive capabilities of its own. To defend against this force, NATO would require organized conventional defenses drawing on contributions from the United States and allied nations. Whether NATO's defense effort would be mounted in Eastern Europe is an open issue, but the Rome Summit made clear that the alliance will continue to protect its borders. Defense of the Oder-Neisse border area of Germany thus seems likely to remain a bedrock focus of NATO's future conventional defense plans in Central Europe. For this reason, the employment concepts outlined by this study seem likely to remain valid for the future, and even greater emphasis will probably be placed on nonlinear forms of ground warfare.

For a detailed appraisal of NATO defense issues in the post-Soviet era, see my report *The Future U.S. Military Presence in Europe: Forces and Requirements for the Post-Cold War Era*, R-4194-EUCOM/NA, 1992.

BIBLIOGRAPHY

Asmus, Ronald A., *The Politics of Modernizing Short-Range Nuclear Forces in West Germany*, RAND, R-3846-AF, 1989.

Biddle, Steven, "The European Conventional Balance: A Reinterpretation of the Debate," *Survival*, March–April 1988.

Bradley, Omar N., *A Soldier's Story*, Rand McNally, New York, 1951.

Bulletin of the Atomic Scientists, "A New European Defense," September 1988.

Chailand, Gerard, and Jean-Pierre Rageau, *Strategic Atlas: A Comparative Geopolitics of the World's Powers*, Harper and Row, New York, 1985.

Chandler, David G., *The Campaigns of Napoleon*, Macmillan, New York, 1978.

Cheney, Dick, Secretary of Defense's Annual Report to the President and the Congress, Washington, D.C., 1990.

Congressional Budget Office, "Budgetary and Military Effects of a Treaty Limiting Conventional Forces in Europe," U.S. Congress, Washington, D.C., 1990.

Congressional Budget Office, "U.S. Ground Forces and the Conventional Balance in Europe," U.S. Congress, Washington, D.C., 1988.

Cordesman, Anthony H., *NATO's Central Region Forces: Capabilities, Challenges, Concepts*, RUSI, London, 1988.

Cuthbertson, Jan, and Peter Volten (eds.), *The Guns Fall Silent: The End of the Cold War and Future of Conventional Disarmament*, Institute for East-West Security Studies, Occasional Paper 17, New York, 1990.

Davis, Paul K., *Planning for Long-Term Security in Central Europe: Implications of the New Strategic Environment*, RAND, R-3977-USDP, August 1990.

Davis, Paul K., *The Role of Uncertainty in Assessing the NATO/Pact Central Region Balance*, RAND, N-2839-RC, 1988a.

Davis, Paul K., *Toward a Conceptual Framework for Operational Arms Control in Europe's Central Region*, RAND, R-3704-NDRI, 1988b.

Davis, Paul K., et al., *Variables Affecting Central-Region Stability: The "Operational Minimum" and Other Issues at Low Force Levels*, RAND, N-2976-USDP, 1989.

Deitchman, Seymour J., *Military Power and the Advance of Technology: General Purpose Military Forces for the 1980s and Beyond*, Westview Press, Boulder, Colorado, 1983.

Department of the Army, *Soviet Army Operations*, Washington, D.C., 1978.

Department of Defense, *Center Region Military Balance Study, 1978–1984*, Washington D.C., 1984.

Department of Defense, Secretary of Defense Annual Reports to Congress, Washington, D.C., 1976–1989.

Department of Defense, *Soviet Military Power: Prospects for Change, 1989*, Washington, D.C., 1989.

Eisenhower, David, *Eisenhower: At War 1943–1945*, Little, Brown, Boston, 1986.

Epstein, Joshua M., *Conventional Force Reductions: A Dynamic Assessment*, The Brookings Institution, Washington D.C., 1990.

Epstein, Joshua M., *Strategy and Force Planning: The Case of the Persian Gulf*, The Brookings Institution, Washington, D.C., 1987.

Erickson, John, et al., *Soviet Ground Forces: An Operational Assessment*, Westview Press, Boulder, Colorado, 1986.

Federal Republic of Germany Ministry of Defense, *NATO's Operational Minimum and Force Buildup of the Warsaw Pact and NATO*, Bonn, 1988.

Fuller, J. F. C., *A Military History of the Western World*, Funk and Wagnalls, New York, 1966.

Gelman, Harry, *The Soviet Turn Toward Conventional Force Reduction: The Internal Struggle and the Variables at Play*, RAND, R-3876-AF, 1989.

Golden, James, et al., *NATO at Forty: Change, Continuity, and Prospects*, Westview Press, Golden, Colorado, 1989.

Griffith, Paddy, *Forward into Battle: Fighting Tactics from Waterloo to Vietnam*, Antony Bird Publications, Sussex, 1981.

Herbert, Paul H., "Deciding What Has to be Done: General William E. Dupuy and the 1976 Edition of FM 100–5, Operations," Leavenworth Papers, Number 16, Combat Studies Institute, U.S. Army Command and General Staff College, Ft. Leavenworth, Kansas, 1988.

Hughes, Wayne P., *Fleet Tactics: Theory and Practice*, U.S. Naval Institute, Annapolis, Maryland, 1986.

International Institute for Strategic Studies, *The Military Balance 1989–1990*, Brassey's, London, 1989.

Joint Staff (U.S.), *Military Posture FY1988*, Washington, D.C., 1988.

Joint Staff (U.S.), *1989 Joint Military Net Assessment*, Washington, D.C., 1989.

Jones, Archer, *The Art of War in the Western World*, University of Illinois Press, Chicago, 1987.

Karber, Phillip, and A. Grant Whitley, "The Operational Realm," in Golden et al., 1989.

Karnowski, Stanley, *The German Army in NATO's Strategy*, National Defense University, Ft. Lesley J. McNair, Washington, D.C., 1984.

Kaufmann, William W., *Glasnost, Perestroika, and U.S. Defense Spending*, The Brookings Institution, Washington, D.C., 1990.

Keeney, Ralph L., and Howard Raiffa, *Decisions With Multiple Objectives: Preferences and Value Tradeoffs*, John Wiley and Sons, New York, 1976.

Kennedy, Thomas J., Jr., *National Politico-Military Consultation: Shaping Alliance Decisions*, National Defense University, Ft. Lesley J. McNair, Washington, D.C., 1984.

Koldoziej, Edward A., and Patrick M. Morgan, *Security and Arms Control, Vol. 2: A Guide to International Policymaking*, Greenwood Press, New York, 1987.

Kugler, Richard L., *The Great Debate: NATO's Evolution in the 1960s*, RAND, N-3252-FF/RC, 1991.

Kugler, Richard L., "On Land Warfare: NATO's Quest for Vision, 1962–1987," in Koldoziej and Morgan, 1987.

Kugler, Richard L., *Laying the Foundations: NATO's Evolution in the 1950s*, RAND, N-3105-FF/RC, June 1990a.

Kugler, Richard L., *NATO's Future Role in Europe: Toward a More Political Alliance in a Stable '1.5 Bloc' System*, RAND, R-3923-FF, May, 1990b.

Kugler, Richard L., "Surviving Forty: Toward a Robust NATO Conventional Defense Posture" in James R. Golden et al. *NATO at Forty: Change, Continuity, and Prospects*, Westview Press, Boulder, Colorado, 1989.

Lawrence, Richard, and Jeffrey Record, *U.S. Force Structure in NATO: An Alternative*, The Brookings Institution, Washington, D.C., 1976.

Levin, Carl (Sen.), "Beyond the Bean Count: Realistically Assessing the Conventional Military Balance in Europe," Senate Armed Services Committee, Washington, D.C., 1988.

Liddell Hart, B.H., *Strategy*, Signet, New York, 1967.

Lupfer, Timothy T., *The Dynamics of Doctrine: The Changes in German Tactical Doctrine During the First World War*, Combat Studies Institute, U.S. Army Command and General Staff College, Fort Leavenworth, Kansas, 1981.

MacDonald, Charles R., *A Time for Trumpets: The Untold Story of the Battle of the Bulge*, William Morrow and Co., New York, 1985.

Mako, William P., *U.S. Ground Forces and the Defense of Central Europe*, The Brookings Institution, Washington, D.C., 1983.

McPherson, James M., *The Battle Cry of Freedom: The Civil War Era*, Oxford University Press, Oxford, 1988.

Mearsheimer, John J., *Conventional Deterrence*, Cornell University Press, Ithaca, New York, 1983.

Natkiel, Richard, *Atlas of 20th Century Warfare*, Gallery Books, New York, 1982.

NATO Information Service, *The North Atlantic Treaty Organization: Facts and Figures*, Brussels, Belgium, 1984.

Pierre, Andrew, et al., *The Conventional Defense of Europe: New Technologies and New Strategies*, Council on Foreign Relations, New York, 1986.

Posen, Barry R., *The Sources of Military Doctrine: France, Britain, and Germany Between the World Wars*, Cornell University Press, Ithaca, New York, 1984.

Presseservice der SPD, *European Security 2000: A Comprehensive Concept for European Security From a Social-Democratic Point of View*, SPD Headquarters, Bonn, Germany, 1989.

Quade, E.S., *Analysis for Public Decisions*, North-Holland, New York, 1962.

Rohn, Laurinda, *Conventional Forces in Europe: A New Approach to the Balance, Stability, and Arms Control*, RAND, R-3732-USDP/AF, May 1990.

Seaton, Albert, *The Russo-German War, 1941–1945*, Praeger, New York, 1971.

Simpkin, Richard, *Race to the Swift: Thoughts on Twenty-First Century Warfare*, Brassey's, London, 1985.

Steinbrunner, John D., and Leon V. Sigal, *Alliance Security: NATO and the No-First-Use Question*, The Brookings Institution, Washington, D.C., 1983.

Stromseth, Jane E., *The Origins of Flexible Response: NATO's Debate Over Strategy in the 1960s*, St. Martin's Press, New York, 1988.

Taylor, James G., *Force on Force Attrition Modelling*, Military Applications Section, Operations Research Society of America, Arlington, Virginia, 1981.

Thomson, David, *Europe Since Napoleon*, Alfred Knopf, New York, 1964.

Thomson, James A. *An Unfavorable Situation: NATO and the Conventional Balance*, RAND, N-2842-FF/RC, 1988.

U.S. Army, Headquarters, FM 100–5, "Operations," Washington, D.C., 1986.

U.S. Army Concepts Analysis Agency, "Weapons Effectiveness Indices/ Weighted Unit Values III (WEI/WUV III)," Bethesda, Maryland, 1978.

U.S. Army Training and Doctrine Command, "Future Warfighting Concepts," Ft. Monroe, Virginia, April 1990.

von Bulow, Andreas, "Conventional Stability, NATO-WTO: An Overall Concept," paper presented at Hearings before the U.S. House of Representatives, Armed Services Committee, Washington, D.C., October 7, 1988.

von Manstein, Erich, *Lost Victories*, Presidio Press, Novato, California, 1984.

von Mellenthin, F. W., *Panzer Battles*, Futura, London, 1979.

von Mellenthin, F. W., and R.H.S. Stolfi (with E. Stobik), *NATO Under Attack*, Duke University, Durham, North Carolina, 1984.

Werth, Alexander, *Russia at War, 1941–1945*, Dutton, New York, 1964.

Wolfe, Thomas, *Soviet Power and Europe, 1945–1970*, Johns Hopkins University Press, Baltimore, Maryland, 1970.