Simulation Support of Large-Scale Exercises

A REFORGER Case Study

Patrick D. Allen

RAND | ARROYO CENTER

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

ISBN: 0-8330-1270-3

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

R-4156-A

Simulation Support of Large-Scale Exercises

A REFORGER Case Study

Patrick D. Allen

Prepared for the
United States Army

RAND

PREFACE

This report documents the results of a research project on "Future REFORGERs," part of a larger effort on Unit Training Strategies conducted for the U.S. Army Training and Doctrine Command and U.S. Army, Europe. Although the initial focus of the project was on REFORGER exercises and their likely forms in the future, the scope of the project expanded to provide insights into the design and conduct of large-scale multiechelon exercises.

This research was conducted under the RAND Arroyo Center's Manpower and Training program and was carried out in part through the Arroyo Center's European office in Heidelberg, Germany.

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

SUMMARY

BACKGROUND

Traditionally, the Army has favored the large-scale, multiechelon exercise because it provided the closest approximation to conditions encountered in actual warfare. But exercises that depend on large numbers of combat vehicles maneuvering freely over a wide area may no longer be a viable training strategy outside of a military installation, or even within most installations. Such exercises are becoming increasingly difficult because of cost, environmental, and political constraints. In addition, as the Army faces a future in which its mission is likely to shift from forward-deployed defense of overseas areas to contingency operations in any part of the world, traditional large-scale ground maneuvers may become even more difficult to perform.

These issues have been particularly important in Germany, where one of the more prominent large exercises, REFORGER (Return of Forces to Germany), took place annually during the 1970s and 1980s. This type of exercise faced growing constraints arising from its increasing cost and the German public's loss of patience with the damage and disruption that inevitably accompany such maneuvers. As a result, U.S. Army, Europe decided to experiment with different ways of conducting exercises to determine if there was a better alternative or combination of alternatives. It selected two exercises, Caravan Guard 89 (CG 89) and Centurion Shield 90 (CS 90—a REFORGER exercise), as vehicles to test alternative exercise modes, including the use of simulations. RAND assisted with the test, and this document reports the results of that effort.

OBJECTIVE

This report has three objectives. First, it describes the various exercise modes used in the CG 89 and CS 90 exercises. Second, it analyzes those exercises to determine the success of each of the various exercise modes and identifies specific problems. Third, it makes recommendations about the design of future exercises.

EXERCISE TRAINING MODES EMPLOYED IN CARAVAN GUARD 89 AND CENTURION SHIELD 90

Both exercises employed four different exercise modes, although not in the same proportion or manner. The modes employed were the following:

- **Field Training Exercise (FTX).** A unit deploys its full amount of equipment and maneuvers across the terrain against a similarly equipped opponent.
- **Command Field Exercise (CFX).** A reduced number of unit vehicles (one per platoon or company) represents the larger unit.
- **Command Post Exercise (CPX).** Only the command elements of a unit are trained, either in the field or in their garrison location.
- **Computer Assisted Exercise (CAX).** Any exercise that employs a computer as one of its assessment tools.

CG 89 featured two opposing divisions with their subordinate elements operating in either the FTX or CFX mode. This "maneuver box" was supplemented with the Corps Battle Simulation (CBS) providing a brigade level flank, and a second simulation, the Warrior Preparation Center's Ground Warfare Simulation (WPC's GRWSIM), filling in the other flank and the rear of each force. CS 90, which took place four months after CG 89 and drew a number of lessons from it, also employed units in the FTX and CFX mode, but at the corps rather than the division level. It employed the same two simulations to fill in the flanks and rears. However, CS 90 also employed German units in the CBS and live portions of the exercise.

FUTURE FOCUS ON SIMULATIONS

Analysis of the two exercises revealed both benefits and problems. The general assessment of CG 89 was that some modes, such as simulations, provided better training for some of the functional areas, such as the intelligence and deep battle elements.[1] However, the "seams" between the various exercise modes were all too apparent, and they detracted from the realism of the exercise. Evaluation of CS 90 led to the conclusion that meshing the various exercise modes was too complex a task. FTXs appear to work best for training at the

[1]"Functional area" is a broad term applied in different ways, depending upon the circumstances. The term does include the Army's standard battlefield operating systems (BOS), but it also includes other areas. Please see the list of functional areas, p. xxiii.

battalion level and below.[2] Above battalion, CAX provides the best opportunity for training all the functional areas simultaneously. This conclusion, coupled with snowballing maneuver costs and restrictions, led us to recommend simulations as the primary training mode, with selected command elements in the field to achieve specific objectives.

OBSERVATIONS REGARDING EXERCISE DESIGN RESULTING FROM THE ANALYSIS OF CARAVAN GUARD 89 AND CENTURION SHIELD 90

Our analysis of the two exercises yielded four primary suggestions for future exercises. These address the training objectives of the exercise, choice of the training mode, ensuring a match between the mode selected and the training objectives of the various functional areas, and a balancing of competing requirements.

Selecting and Articulating Training Objectives

Typically, a unit's Mission Essential Task List (METL) defines the tasks viewed as the most important to train. The METL works well at levels below the division, but defining tasks for corps and above poses some difficulties, particularly defining a full METL. For exercises, we suggest that if a full METL is not yet prepared, the commander should develop one for each training element (e.g., maneuver element, staff G1–G5, deep battle cell) in somewhat less detail than a fully developed METL but made more complete than the broad exercise objectives.

Once defined, the objectives should be clearly articulated. A lack of adequate detail in the training objectives affects exercise design, execution, and evaluation. Further, training objectives should be prioritized to allow resolution of conflicts between different functional areas. For example, accomplishing an operational objective may require an unrealistic movement of a unit; this move might work to the detriment of an intelligence training objective. An established priority will speed the resolution of the conflict between objectives. For example, the priority may be stated as follows: "In case of conflicting objectives, maneuver training objective number six (execute a flank counterattack) will outweigh intelligence training objective number four (accurately predict arrival time of follow-on units)."

[2]A battalion-level FTX often includes a brigade headquarters and a "slice" of brigade support. This recommendation does not discourage this practice.

Selecting the Training Mode

Each training mode offers different advantages and exacts different penalties. Training objectives provide the key to the selection of the mode. The experience of CG 89 and CS 90 leads us to favor a single training mode for any given exercise. Training modes are not always compatible, and problems with the interfaces between modes lead to a number of problems in exercise design and execution. For example, the reduced number of vehicles in the CFX mode poses real challenges for intelligence staffs, since most of the visual cues they use to locate and identify units are no longer present. Similarly, when live units are being faced by simulated units, the visual cues live units use to locate and engage enemy forces are not present.

From these and many other examples, this report concludes that when simulations are the primary training mode, there should be no maneuver units in the field. Trying to use both simulations and units in the field in the same exercise tends to be expensive, and it detracts from achieving the training objectives.

We do not recommend that *every* exercise use simulation only. There are many reasons why some exercises should include units in the field, such as deployment or logistics exercises. Moreover, there is an advantage to trying out everything in the field to make sure that all elements work well together and that the simulations are not missing something. But given the large number of difficult problems that arise when multiple training modes are used together, we recommend that whenever possible, a given exercise should use only a single training mode.

Testing an Exercise Design

With objectives defined and the mode selected, the exercise design should be tested from two aspects prior to the exercise. First, pretest the exercise design to ensure that it demands the *focus of attention* of each training audience functional area on the training objectives. Second, pretest the exercise design to ensure that the *sequence of events* provides the opportunity for the training audience within each functional area to accomplish its training objectives. Although both components may be tested simultaneously, the first tends to be time independent and the second is time dependent. The first component focuses on the organization of the training audience, scenario, and training modes. The second component focuses on the sequence of events desired, from the perspective of each training element. Pretesting does not require execution of the full exercise. It can be done by someone not involved with the design of the exercise, who will review the design with an eye to all functional areas to identify any problems between them.

Striking a Balance

Exercise designers must strike a balance among a number of competing considerations, but two issues are particularly important. One is the choice of the degree of freeplay and scripting in an exercise; the other is to make clear the distinctions between the training audience and training support personnel. Rigidly scripted exercises generally ensure that all specific training objectives get accomplished, but they stifle initiative and some elements of realism because the players cannot affect the course of events. Also, they tend not to be very good at presenting the surprising situations that are more typical of actual conflict.

Freeplay exercises encourage initiative and present unexpected events, but they are more difficult to control and do not guarantee accomplishment of any of the specific training objectives. For example, if one training objective is to train for a reconstitution operation, and none of the forces are significantly attrited in a freeplay exercise, then that training objective will not be accomplished. Although a freeplay exercise will tend to accomplish many of the broad training objectives, a freeplay exercise cannot guarantee the accomplishment of many specific training objectives. This is why the balance between the freeplay and scripted components of an exercise is so important.

The balance between the training audience and support personnel depends on many things, including the mode selected. All exercises require support personnel, and some modes require more than others. The exercise designers must decide how many resources are available to support the exercise. The training objectives should be the primary determinant. Whatever the choice, designers should resist the temptation to include the support staff as part of the training audience. The response cells supporting the exercise are there to buffer or protect the training audience from the artificialities of the exercise. If they view themselves as part of the training audience, they tend to try to win the battle by taking advantage of their knowledge of the game.

For example, personnel placed in front of a computer screen and told they are part of the training audience will play to win, and will tend to figure out how to beat the machine at its own game. This is called the "videogame syndrome." If these personnel are supposed to be providing training support to the training audience, they are supposed to buffer the training audience from the artificialities of the game, rather than exploit those artificialities to win the game.

SUGGESTIONS FOR FUTURE EXERCISE DESIGN

The analysis of the two exercises leads to three major recommendations for future large-scale, multiechelon exercises.

- First, given all the constraints faced and the increasing costs, most future large-scale exercises should consist of a single training mode, and the preferred training mode should be full simulation, as opposed to a combination of simulation and field exercises.[3]
- Second, if simulations become the primary mode for large-scale exercises, a number of limitations affecting the current family of simulations must be overcome.
- Third, whenever possible, exercises should include both active and reserve component units and forces from other services and nations.

Full Simulations for Future Exercises

Results of the two test exercises clearly show that mixing modes causes considerable problems. Thus, we recommend the selection of a single training mode for a given exercise.[4] Further, because of the increasing constraints on maneuvering forces in field, the operational costs of using live heavy forces, and the recent advances in combat simulation technology, we recommend that simulations be used as the primary training mode for large-scale, multiechelon exercises. The costs and problems associated with FTXs argue strongly against that choice, at least for large exercises. Also, a larger number of the functional areas training simultaneously derive greater training benefit from using simulations. This does not mean that all exercises should use simulations. A deployment exercise needs to deploy large amounts of equipment, and FTXs work very well at certain levels. But simulations do a better job for many types of training objectives in large-scale, multiechelon exercises.

---

[3]As clarification, we recommend that the training audience operate from headquarters in the field through distributed wargaming, but no maneuver elements should be in the field when a combat simulation is the selected training mode.

[4]If significant technological advances occur that allow simulated and live forces to interact consistently in a seamless manner, then this recommendation will need to be reconsidered.

Simulations Require Improvement

Simulations show considerable promise as a training tool. However, the current suite of combat simulations has a number of problems in the way its simulations represent different functional areas. The broad areas requiring improvement in mid- to high-intensity combat simulations are described below.

- **The representation of combined arms effects.** Different branches of the three primary combat arms (armor, infantry, and artillery) are better suited to certain combat situations than others. A mix of combined arms in each of these categories is essential in most frequently encountered combat situations.[5] For example, terrain multipliers in models that account for differences in topography tend to be applied identically. Thus, a tank company gains the same advantage from rough or urban terrain as a light infantry unit—which would not be the case in reality. Also, most models impose no penalty for having no infantry in a force.
- **The types of battles.** Current models have difficulty representing key features of different types of battles (e.g., river crossings, flank attacks, or counterattacks) or phases of a battle (e.g., breakthrough, exploitation, and pursuit).
- **Aspects of the depiction of the operational level of war.** The simulations do not recognize some of the fundamental differences that occur at different echelons. For example, divisions take longer to move than battalions or companies. But models tend to move all units at the rate of their unit of lowest resolution. Thus, if the model portrays companies, divisions in the model move at nearly the same speed as companies—which is not realistic.
- **The "fog and friction of war."** Simulations tend to ignore the fact that things often go wrong in war—units get lost, the wrong ammunition arrives, or units miss departure times.
- **Intelligence functions and products.** Models frequently provide information that is too good; that is, they pass information on units that is far more detailed than could ever be obtained by a single intelligence asset. In essence, the model provides the players the ground truth, a situation that rarely if ever occurs in combat.

---

[5]Since the CG 89 and CS 90 exercises, the Army has implemented the COBRA addendum to the Corps Battle Simulation (CBS) addressing most combined arms effects.

Correcting the limitations of simulations poses a complex task. The Army agency responsible for acting as the central clearinghouse of identified problems in most Army training simulations is the National Simulation Center. All identified problems are collected and prioritized, and then resources are allocated to correct these problems. The difficulty is that the types of problems that should be considered first-order problems, such as an adequate representation of friction and the difficulties of passage of lines operations, are not usually raised to the top of the list by the user community. When the representatives of the user community focus on tactical issues, the more aggregate operational issues tend to be ignored. One goal of this report is to inform the user community of the types of basic issues that simulations do not yet adequately address, so that they may be raised in priority for earlier resolution.

Expand the Set of Exercise Participants

The anticipated direction of future force structures makes it more important than ever that the active and reserve components train together. Future exercises above brigade should incorporate reserve components into the design. These units could participate in exercises without leaving their home station by means of a distributed wargaming system.

It is equally important to incorporate forces from the other services. Operations Just Cause and Desert Storm both highlighted the importance of interservice procedures. The ability to work smoothly with members from another service will not occur without training. In fact, the Army's doctrine assumes joint participation.

Also, the likelihood of coalition warfare also appears to be increasing, and exercises should include the forces of other countries as well. Such participation would require some way to address inherent doctrinal differences that would have to be addressed in any contingency operation. We recommend that the United States develop a standard doctrine for coalition operations.

ACKNOWLEDGMENTS

The author is grateful for all of the cooperation provided by the USAREUR REFORGER Planning Group. In particular, Major Thomas Hill has been the continuing presence and point of contact for all of our work with the group.

Special thanks to Mr. Tom Polsley, an employee of the U.S. Army Training and Doctrine Command (TRADOC); he was assigned to the Arroyo Center's USAREUR office in 1989–1990. In addition, Major Tony Kitts and Major Howard Lee of the National Simulation Center provided valuable information on the current and projected status of the Corps Battle Simulation and its configuration control process.

Members of the RAND team that participated in the data collection and analysis, as well as the direction of this project, include Michael Polich, Thomas Lippiatt, Leland Pleger, James Kahan, and Robert Worley. A special note of thanks goes to Leland Pleger for handling this project during the difficult early stages of project definition. A number of his original concepts have been included in this final report.

Thanks are also due to the reviewers of this document: William Whelan of RAND and LTC Charles Thompsen, a RAND Army Fellow. Additional assistance was provided by Jerry Sollinger, a RAND communications analyst who contributed significantly to the extensive revision of this document. All opinions and misstatements remain the responsibility of the author.

CONTENTS

FIGURES

TABLES

GLOSSARY

ACE	Allied Command Europe
ARTEP	Army training and evaluation program
ATOC	Allied Tactical Operations Center
AWSIM	Air War Simulation, housed at WPC
BAI	Battlefield air interdiction
BCTP	Battle Command Training Program
BDE	Brigade
BOS	Battlefield operating system
C2	Command-and-control
CAX	Computer assisted exercise
CB	Counterbattery (artillery fires)
CBS	Corps Battle Simulation (a combat model)
CENTAG	(NATO's) Central Army Group
CFX	Command field exercise
CG 89	1989 Caravan Guard exercise
Class I	Supplies—rations
Class II	Supplies—individual equipment
Class III	Supplies—petroleum, oil, and lubricants
Class IV	Supplies—construction materials
Class V	Supplies—ammunition
Class VI	Supplies—personal items
Class VII	Supplies—major end-items
Class VIII	Supplies—medical items
Class IX	Supplies—repair parts
CONUS	Continental United States
CPX	Command post exercise
CS 90	1990 Centurion Shield exercise
CSS	Combat service support
DWS	Distributed wargaming simulation
ECC	Exercise Control Center
EW	Electronic warfare
FEBA	Forward edge of the battle area

FFRDC	Federally funded research and development center
FLOT	Forward line of own troops
FTX	Field training exercise
G1	General Staff section for personnel (and its head)
G2	General Staff section for intelligence (and its head)
G3	General Staff section for operations (and its head)
G4	General Staff section for logistics (and its head)
G5	General Staff section for civilian affairs (and its head)
GRWSIM	Ground War Simulation, housed at WPC
hex	hexagon—the basic geographical unit for many models
ICM	Intelligence Collection Model
JRTC	Joint Readiness Training Center
JSEAD	Joint suppression of enemy air defenses
METL	Mission essential task list
MILES	Multiple integrated laser engagement system
NAI	Named area of interest
NATO	North Atlantic Treaty Organization
NTC	National Training Center, Fort Irwin, California
OPTEMPO	Operational tempo
POMCUS	Prepositioned Overseas Material Configured in Unit Sets
RAOC	Rear area operations center
REFORGER	Return of Forces to Germany exercise
SOF	Special operations forces
TACFIRE	Tactical fire control system
TAI	Target area of interest
UCC	Umpire Control Center of an exercise
USAREUR	United States Army, Europe
UTM	Universal transverse mercator (map coordinates)
WPC	Warrior Preparation Center, Einsiedlerhof, Germany

LIST OF FUNCTIONAL AREAS

Functional Areas	Battlefield Operating Systems
Command-and-control (C2)	Command and Control
Signal	Command and Control
Deception	Command and Control
Electronic Warfare (EW)	Command and Control/ Intelligence
Intelligence (G2)	Intelligence
Maneuver	Maneuver
Operations (G3)	Maneuver
Close operations	Maneuver
Deep operations	Maneuver
Rear operations	Maneuver
Special Operations Forces (SOF), Long-Range Reconnaissance Patrol (LRRP), and Spetsnaz	Maneuver
Airborne/air assault	Maneuver
Army aviation	Maneuver/Fire Support
Air	Fire Support
Fire support (non-air)	Fire Support
Joint Suppression of Enemy Air Defenses (JSEAD)	Fire Support/Air Defense
Air defense	Air Defense
Engineers	Mobility, Countermobility, and Survivability
Combat Service Support (G4)	Combat Service Support
Personnel (G1)	Combat Service Support
	Other Categories
Corps picture	Exercise design issue
Threat play	Exercise design issue
Attrition rates	Simulation issue
Combined arms	Simulation issue
Unit resolution	Simulation issue

1. INTRODUCTION

BACKGROUND

Traditionally, the Army has favored large-scale, multiechelon field exercises. Such exercises involve the maneuvering of substantial numbers of combat units and their support elements against a similarly equipped opponent. REFORGER (Return of Forces to Germany) is probably the best-known example, but other exercises, such as Caravan Guard conducted by V Corps in Germany, share many characteristics with REFORGER.

Purposes of Large-Scale Exercises

A number of reasons favor large-scale (above brigade) exercises, but three stand out: they provide the best approximations to actual combat (realism), they use resources efficiently, and they reflect national capability and commitment. Army training strives to replicate actual combat as closely as possible. The exercises at the National Training Center (NTC) and the Joint Readiness Training Center (JRTC) are the most successful in this regard, but they involve at most a brigade. The large-scale, multiechelon exercise offers the best approximation of combat for a sizable force, such as a division or greater.

This type of exercise allows the Army to see how well all the components of a force work together. This observation applies to the parts both within a given echelon and, equally importantly, between echelons. Commanders also value the large-scale field exercises because they more closely approximate the "fog and friction" found in actual combat. It is impossible to anticipate all the unexpected situations that occur in war. Large-scale exercises provide the same sorts of events and thus hone the skills of the participants in dealing with them.

Yet another advantage of the large exercise is that it helps reduce the amount of resources necessary to train different echelons and functional areas.[1] To quote FM 25-100:

---

[1]"Functional area" is a broad term applied in different ways, depending upon the circumstances. The term does include the Army's standard battlefield operating systems (BOS), but it includes other areas as well. Please see the list of functional areas presented on p. xxiii.

> Limited time and other resources do not permit developing sequential training programs, in which each echelon from lower to higher is successively trained to reach interim "peaks" in proficiency. Therefore, leaders use a multiechelon training approach to plan training events. Multiechelon training allows simultaneous training and evaluation on any combination of individual and collective tasks at more than one echelon. Multiechelon training is the most efficient and effective way of training and sustaining a diverse number of mission-essential tasks within limited periods of training time.[2]

Another motivation for large exercises is the fact that they demonstrate both capability and political will. REFORGER again provides the best-known example. The annual deployment of a large force to Germany showed the Warsaw Pact that the United States was committed to meet its treaty obligations.

Large-Scale, Multiechelon Exercises Face Growing Constraints

Despite their benefits, traditional exercises are facing increasing constraints in two main areas: decreasing resources and growing environmental concerns. First, field exercises have become expensive, in both operating costs and maneuver damage costs. Modern weapon systems, such as the Apache and M-1, that depend on high-technology components simply cost more to operate. Also, inflation and the growing urbanization of Germany have increased the costs of maneuver damage. Examples of these costs appear in Section 2.

In addition to the increased expense, growing concerns over environmental damage are also working to limit the scope and frequency of maneuvers. The German populace is growing less willing to tolerate the noise pollution, traffic jams, fuel spills, and physical damage that inevitably accompany large exercises. And it is becoming increasingly vocal in letting local politicians know of its dissatisfaction with this environmental damage.

Of course, the German public's increasing dissatisfaction with the large exercise stems in part from the dramatic change in the geopolitical situation, which has undermined a large portion of the rationale for these exercises. Now that the Warsaw Pact has dissolved and the former Soviet Army is withdrawing behind its own borders, and the Soviet Union has reconstituted as a loose confederation of republics, the German public feels far less threatened. Thus there is less ap-

[2]FM 25-100, p. 3-19.

parent need to demonstrate resolve through large troop deployments, further lowering the public's tolerance of the inconvenience and damage that the exercises cause.

In light of these increasing costs and constraints, commanders question whether the REFORGER objectives still apply and whether they are still being met. The numerous restrictions placed on the movement of combat vehicles have sharply curtailed exercise realism. The inability to have free maneuver play legitimately calls training value into question. Also, there is less need now to physically demonstrate the will to reinforce NATO.

These constraints and questions led the Army to test alternative exercise modes. The exercises selected for the test were Caravan Guard 89 (CG 89) and the 1990 REFORGER exercise, called Centurion Shield (CS 90). These exercises were to experiment with a variety of exercise modes—both field exercises and simulations—and determine which combination worked best for accomplishing the training objectives under the limitations described above. RAND was asked to assist in the evaluation of the experimental exercise designs and executions and to make suggestions on the design of future exercises.

PURPOSE

This report has three purposes. First, it analyzes the design and execution of the CG 89 and CS 90 exercises to determine how well the various training modes worked and what problems surfaced.[3] Next, it draws lessons from that analysis and, finally, makes some recommendations about future exercises and the simulations to support them.

ORGANIZATION

Section 2 discusses experimental field exercises in general, highlighting the various modes considered and discussing some of the inherent tradeoffs of the various modes. It then analyzes the exercise designs of CG 89 and CG 90. Section 3 presents the four major lessons drawn from the analysis. Section 4 contains our recommendations for future exercises.

---

[3]These exercises receive more detailed treatment in Allen et al. (1992a, 1992b).

2. TYPES OF EXPERIMENTAL EXERCISES

This section discusses the general types of training modes available, how they were employed during the CG 89 and CS 90 experimental exercises, and the benefits and problems associated with these exercises.

TRAINING MODES

The commander of each unit is responsible for training that unit to be ready to accomplish its assigned missions. The commander has a number of training modes from which to select, and combinations of training modes are usually required to successfully train a unit in all of its essential tasks. The commander's choice of which training mode or modes to employ in a given exercise depends primarily on the training objectives, subject to such constraints as time, cost, facilities, supporting manpower, and other resources.

Commanders know that every training mode lacks the realism of combat to some degree, but combat is an expensive and unforgiving mode of training. Peacetime exercises are designed to provide some of the aspects of combat at acceptable costs, with the opportunity for repetition if needed.

Each mode facilitates different aspects of training. The commander's task, then, becomes one of sorting among the various modes available to select the ones that best accomplish the specific training objectives. FM 25-100 lists thirteen types of training events.[1] This report will describe only four, since these were the primary training modes used during CG 89 and CS 90: FTX, CFX, CPX, and a variation on CPX, CAX.

The Field Training Exercise (FTX)

An FTX consists of full units maneuvering with all of their assigned equipment and personnel in the field. Units may maneuver anywhere within the terrain box subject to predefined maneuver restrictions. Independent umpires, controllers, or evaluators observe engagements, determine results, and assess combat outcomes.

---

[1]FM 25-100, p. 3-12.

FTXs have the major advantage of offering the most realistic training. Maneuvering full units over actual terrain to confront a live opponent causes commanders and their staffs to encounter many of the problems that actual combat presents. Further, this type of training will exercise the many interfaces between echelons of units. A major disadvantage of an FTX is its dollar cost in terms of operating tempo and maneuver damage. More recently, exercise restrictions, designed primarily to reduce maneuver damage costs, reduce the FTX's realism as well.[2]

FTX Costs

Real dollar costs generally stem from two sources: operating costs and maneuver damage costs.

Operating Costs. These are normally expressed as OPTEMPO costs, which stands for "operational tempo" and refers to the cost associated with sustaining a level of training activity. Similar to "flying hours" for the Air Force or "steaming hours" for the Navy, OPTEMPO is defined for the Army as the cost of the fuel, spare parts, and related items accrued for each mile driven by a particular type of vehicle. The examples in Table 1 are taken directly from FM 25-100, p. 3-14.

Based upon USAREUR historical data, the OPTEMPO costs for exercising an M1 tank maneuver battalion with all of its vehicles in the field for a two-week exercise were about $984,000. An M3 (armored cavalry) maneuver battalion for the same period would cost about $249,000.[3]

Maneuver Damage Costs. Based upon an agreement between the United States and Germany, the United States pays 75 percent of maneuver damage costs and Germany pays 25 percent. The annual average cost for the U.S. portion of maneuver damage for all exercises in Germany, including REFORGER, is about $33 million. Table 2 lists the U.S. portion of REFORGER exercise maneuver damage from 1980 through 1988. The decrease in maneuver damage costs since 1984 stems from greater maneuver restrictions and from the fact that not all of the claims against the more recent exercises have been

[2]One problem with a multiechelon FTX is that as the troops perceive it, the higher the echelon of the exercise, the more of their time is wasted in waiting and generally being used as training aids for the higher commanders.

[3]By comparison, a light infantry battalion two-week exercise would cost between $50,000 and $90,000, with the cost varying as a function of the weather (cold weather brings higher costs) and the number of batteries required for radios and secure gear.

Table 1

Sample OPTEMPO Cost Calculations

System	Number Used		Miles Traveled		Cost Factors ($)[a] Class IX		Class III		System Cost ($000s)
M1 Tank	58	×	70	×	(42.00	+	6.40)	=	196.5
M3 CFV	6	×	85	×	(9.25	+	1.05)	=	5.2

[a]For definitions of the cost factors, see the Glossary.

Table 2

U.S. Costs of REFORGER Maneuver Damage, 1980–1988

Year	Cost ($ Millions)
1980	10.7
1981	10.0
1982	9.8
1983	14.3
1984	19.0
1985	8.7[a]
1986	3.1
1987	4.9
1988	4.2
Total	84.7
Average cost per year	9.4

[a]Maneuver restrictions imposed.

settled. USAREUR is still paying claims for maneuver damage from an exercise that occurred ten years ago.[4]

Tracked vehicles tend to cause more maneuver damage than wheeled vehicles. Their weight and cross-country capability provide ample opportunity to destroy fields, crops, and walls, and if they are driven on highways they often damage the road surface. Even heavy wheeled vehicles can cause significant damage, in collisions and cargo spills. In one case, thousands of gallons of fuel spilled at a refueling site. German law has strict standards for cleanup at a fuel spill, mak-

[4]Since maneuver damage costs are paid in German marks, their decrease is even more significant than the dollar amounts alone would indicate. Maneuver damage costs have been decreasing even though the cost of an equivalent claim has risen because of the mark's growth in value relative to the dollar.

ing such accidents very expensive. Generally, more vehicles in the field increases the likelihood of maneuver damage, and the heavier the vehicles, the higher the overall cost of the damage per incident.

Maneuver Restrictions. In early REFORGER exercises, there were few restrictions on where troops and vehicles could move within the exercise maneuver box. Because of this, participating personnel at the tactical level felt they could fully utilize the terrain, and therefore perceived that they had received good, realistic training.

However, this same "good" training caused damage, so maneuver restrictions were imposed. For example, tracked vehicles were required to stay on roads until they reached their position for deployment. Cross-country maneuvers were severely restricted unless the enemy was engaged. This did decrease the maneuver damage costs, as Table 2 shows. At the same time, it significantly reduced the realism of the training for heavy tracked vehicles. Additional maneuver restrictions are being imposed every year.

The Command Field Exercise (CFX)

A CFX consists of command vehicles only, such as the tank platoon leader or even tank company leader, so that the number of personnel and vehicles in the field is reduced. This training mode reduces the costs associated with maneuver damage, operations, and personnel. It also reduces the time that noncommand personnel must wait for something to happen. However, in this mode of exercise an umpire has a much more difficult time of assessing combat and maneuver outcomes, since he must judge whether the tactics of each side could have been implemented as described by the CFX unit commander; in the FTX mode, the umpire can observe the actual implementation.

CFX Costs

Different echelons may be selected as the level of resolution in the CFX exercise. For example, a unit in CFX mode may use command vehicles down to the platoon level. Similarly, one may employ command vehicles down to only the company or even the battalion level, depending upon the training objectives. Based upon USAREUR historical data, the OPTEMPO costs for exercising an M1 (tank) maneuver battalion down to platoon command vehicles in the field for a two-week exercise are about $286,000. An M3 (armored cavalry) maneuver battalion for the same period in the same mode would cost about $110,000.

Assuming that twelve M1 tank battalions and eight M3 battalions were involved in CS 90, the OPTEMPO cost for command vehicles was estimated to be only $4.31 million. Had all of these battalions been fielded in the FTX mode, the cost would have been $13.8 million. The reduced OPTEMPO avoided costs of $9.5 million, while maneuver damage costs declined owing to fewer tracked vehicles in the field. To further reduce the maneuver damage costs, military wheeled vehicles were substituted for many tracked vehicles. The total costs of the CS 90 exercise have not been determined, but the estimates presented above indicate that these measures significantly reduced both the OPTEMPO costs and maneuver damage costs.

The Command Post Exercise (CPX)

Only the command and staff elements of the participating units, such as battalions and above, are the training audience in a CPX. This method is even less expensive in terms of maneuver damage, operating, and personnel costs. However, there are also tradeoffs associated with a CPX's training benefit.

The "realism" of a CPX depends upon both the exercise design and the apparent realism of the training mode selected to assess outcomes of actions initiated by the training audience, such as maneuver, combat, and logistics. The disadvantage of CPXs is that they tend to be less realistic: the friction or problems that could occur below the training audience are either assumed away or scripted in by the exercise directing staff (di-staff). In CPXs, units rarely get lost, become significantly delayed through internal problems, or get sent the wrong types of supplies, unless such events are explicitly inserted by the directing staff.

There are many different ways to assess the events in a CPX. One way is to use referees or controllers, similar to umpires in the field, who look at a situation and determine the outcome. Another method is a manual wargame that employs "look-up" tables to determine the results of a given situation. A third method is to employ a computer as the assessment tool, which is called a CAX (described below).

CPX Costs

CPX costs are normally limited to the support personnel required and modest expenditures for maps and overlays to keep track of the situation. As a result, CPXs tend to be the least expensive training mode examined in this study.

The Computer Assisted Exercise (CAX)

An exercise that employs a computer as one of its assessment tools is, by definition, a CAX.[5] A CPX that employs a computer to assist in the assessment of outcomes is considered a CAX. CAXs have a number of advantages and disadvantages related to realism, manpower requirements, and costs.

Compared with traditional training methods, CAX may provide a better opportunity to train certain functional areas. For example, the G2 training provided by a "scripted" exercise tends to be relatively poor.[6] In a rigidly scripted exercise the events are predefined to unfold in a particular manner, independent of the training audience members' decisions or actions. As a result, the G2 staff tends to support the training of the G3 and the maneuver elements but to not train in their own tasks. Due to the flexibility and uncertainty portrayed in a more freeplay CAX, the G2 staff can receive training in their area of responsibility. During CG 89, the deep operations cell (consisting primarily of G2 personnel but also with personnel assigned from other staff functions) was presented with a deep enemy threat that they had to track and plan interdiction operations against under heavy time pressure. This was better training than they had received through most other training modes.

But current combat training simulations have a number of limitations that need to be overcome before they can become the primary training mode in large-scale Army exercises. For example, during the CG 89 exercise, the deep operations cell was predicting the movement of a simulated enemy division approaching the battle area. Using threat doctrinal movement rates, the cell predicted the progress of the enemy division. Because of aggregation problems in the simulation, however, the threat division was moved at twice the doctrinal movement rate sustained over 24 hours. This lack of realism reduced the training benefit to the deep operations cell. The current limitations of combat simulations that affect training are addressed in detail in Section 4.

---

[5]Sometimes, computers are used to assist in the assessment of predominantly field exercises such as those at the National Training Center (NTC) at Ft. Irwin, California. At the NTC, computers assist in the assessment of indirect fire such as artillery, since the direct fire assessments are handled by the MILES laser system. See Goldsmith, Hodges, and Burn (1990).

[6]In a later example we examine the benefit of using simulations to train G1 and G4 staff elements.

CAX Costs

CAXs are often more expensive than traditional CPXs. This extra cost—surprising to some—results from expenses associated with employing computers, added personnel to operate the computers, and increased communications assets to connect the computers to each other and the rest of the training communications net. It is difficult to determine the costs associated with a CAX, since the costs of software development, hardware purchase, and computer user training are often considered "sunk" costs. In reality, one should amortize these costs over the lifetime of the hardware or software, but this is complicated since the same software can run on many versions of hardware, and the same hardware can use may versions of software. Therefore, one can usually look at only the "marginal" costs of a CAX.

The marginal costs of a CAX include the communications costs and personnel employed only in supporting a specific exercise. For example, the Corps Battle Simulation (CBS) combat model can link the different training sites in the field to the main computer through 9.6 kilobit (KB) bandwidth communications lines. The total communications cost for a three-echelon exercise (brigade through company or division through battalion) is about $30,000. The Warrior Preparation Center (WPC) simulations require a wider bandwidth in order to distribute their games to remote sites. Since the WPC's ground game cannot be distributed quickly enough on a 56 KB line, the next-higher option, a 2000 KB line, must be used. The communications cost associated with the wide bandwidth is just under $100,000 per day, or about a million dollars for a REFORGER exercise. The NATO allies have not enthusiastically endorsed CAX, primarily because of these high communications and personnel costs per exercise. The U.S. Army will be similarly concerned about future CAXs that require hardware and software with high communications and personnel costs.

Personnel Costs of CAX. These are difficult to quantify, since soldiers draw pay whether they are exercising in the field or remaining in garrison. However, two marginal costs are associated with personnel: temporary duty travel and related expenses, and the opportunity costs associated with using personnel who are not part of the training audience (called augmentees or borrowed military manpower). For example, the number of augmentees required to run the simulations at the WPC for CS 90 exceeded 200 people. These training support personnel were taken from their regular assignments to provide support for the primary training audience in the field. Although these costs are difficult to quantify, a good rule of thumb is the more people involved in the exercise, the higher the personnel costs. (See the section on training support personnel in Section 3.)

The CAX Tradeoff. Overall, the increased marginal costs of CAXs and the lack of realism in current simulations create a substantial tradeoff against their increased training benefits. If the costs are reduced (by requiring fewer personnel and less expensive communications) and the quality of the exercises is improved (primarily by improving exercise design and correcting the current limitations in the simulations), then CAX may become the dominant training tool for large-scale exercises.

However, one should not take these required improvements for granted. The limitations in the current combat simulations have been present for years and yet remain unsolved, owing to the way improvements are prioritized. The priority of a given change to fix a problem is based upon inputs from representatives of the user community. They decide what is important, and a fix's priority is defined by its perceived training benefit and the cost and time needed to implement it. The problem is that many users tend to focus on tactical or detailed issues rather than on the more aggregate and operational issues. A detailed discussion of these simulation problems and the prioritization issue is presented in Section 4.

Personnel costs tend to remain high because the simulations used for training tend to be manpower intensive. In these simulations, almost no decision is made automatically by the computer. A person must move every unit, order every attack, and often requisition every supply. Automated and semiautomated command and control needs to be included in the simulations to reduce the high manpower requirements in support of these exercises.

Communications costs remain high because of the wide bandwidth requirements for some combat simulations. The Army may wish to focus on those simulations with lower bandwidths or reduce the bandwidth of larger simulations in order to reduce the communications costs.

Based upon the preceding problems, the advantages of the CAX have not yet convincingly outweighed all of its penalties.

Recent Army Experience with Large-Scale CAX

With the advent of more powerful computers and more sophisticated software, the Army has consistently increased its employment of CAX exercises. CAX exercises range anywhere from a relatively simple program run on a personal computer to more elaborate groups of models run simultaneously on a network of computers. Two examples of frequently employed CAX are the Battle Command Training Program (BCTP) and the Warrior Preparation Center (WPC).

The BCTP warfighter is the first systematic approach to training a higher echelon (either a corps or division headquarters) with a set of measurable standards for specific functional areas. A BCTP warfighter exercise presents a set of situations to the training audience to determine how well it performs certain staff functions. The BCTP consists of two parts: the seminar wargame used to build the team and achieve a common understanding of how doctrine applies to unit tasks, and the warfighter exercise, which is a training event conducted several months after the seminar exercise. The warfighter focuses on diagnosing the unit's proficiency at previously learned tasks, including tasks learned during the seminar exercise. A large number of observer/controller personnel participate in the warfighter exercise to evaluate the training audience. The Corps Battle Simulation (CBS) is the computer wargame used in the BCTP seminar and warfighter exercises.

A BCTP usually focuses on several functional areas within a single echelon. Although the primary emphasis is on G3 and G2 staff procedures, some training is provided for G1 and G4 as well. BCTP presents a "human" and reactive threat to the training audience to instill an essential degree of competition against a thinking opponent. Repetition is allowed only if the training benefit would be greater by repeating previous events rather than continuing from a previous result. BCTP warfighter exercises focus on horizontal synchronization within a single echelon, rather than on vertical synchronization between echelons.[7]

The WPC is a training support facility funded and manned by the U.S. Army Europe and U.S. Air Forces in Europe. The WPC has a large suite of computers and simulations that are used to exercise both air and ground forces in an interactive manner. WPC exercises are also two-sided, with a reactive, thinking opponent or threat cell.

---

[7]The Army also experimented with running two simultaneous BCTP warfighter exercises—one for division and one for corps. Attempts to meet the training objectives of the division competed with attempts to meet those of the corps. As a result of the experiment, the Army adopted as policy that such simultaneous BCTP exercises be the exception rather than the rule. REFORGER 92 plans to run a corps-level BCTP as part of the overall exercise, but the BCTP has priority on accomplishment of the training objectives.

An exercise at the WPC tends to be less scripted and more freeplay, which creates both advantages and disadvantages. The advantage is that some consider completely freeplay exercises more realistic than exercises with any scripted events. The disadvantage is that it is often more difficult to ensure that a freeplay exercise meets specific training objectives. As one might expect, there is a desirable balance between freeplay and scripted exercises, as discussed further in Section 3.

WPC exercises usually involve multiple echelons and many functional areas. It is not unusual to have three or four echelons involved in the training audience, although one may question the level of training provided at the lowest echelon when so many echelons are involved. The reason is that the division blurs between the training audience and the response cells[8] at the lowest echelons in WPC exercises. This issue is discussed in more detail in Section 3. For more information on the WPC and its exercises, see Allen (forthcoming).

These four exercise modes formed the building blocks of the two experimental exercises. Both exercises would retain a sizable FTX, two divisions in one case and two corps in the other. But both exercises would also employ CFX, CPX, and CAX in an attempt to determine whether a different mode or combination of modes would alleviate some of the cost and maneuver restrictions and still provide good training benefit. With this discussion as a backdrop, the section below describes the organization and conduct of the two test exercises.

DESCRIPTION OF CARAVAN GUARD 89 AND CENTURION SHIELD 90

During CG 89 and CS 90, USAREUR experimented with different mixes of training modes. USAREUR wished to determine which provided the best training for each functional area in each echelon and which training modes worked well with each other. If mixing modes proved feasible, effective, and economical, USAREUR would use the lessons learned on these exercises to shape future large-scale exercises in Germany and, perhaps, throughout the Army.

[8]The members of the response cell are not part of the training audience. Rather, they add realism to the exercise by providing information to the training audience in the form it is expected.

Caravan Guard 89: The Training Objectives

The training objectives for CG 89 were the following:

- Test the FTX, CFX, and CPX simulation concept.
- Conduct selected unit ARTEPs.
- Conduct division, brigade, and battalion battle staff training.
- Hone synchronization of the close battle.
- Train to focus combat power.
- Conduct and refine long-range surveillance unit operations.
- Conduct airmobile operations.
- Conduct jamming operations.
- Conduct rear area operations.
- Conduct restricted night operations.

Caravan Guard 89: The Exercise Battlefield

The U.S. V Corps Caravan Guard exercise took place on 13–20 September 1989. A parallel Air Force live-fly exercise called Cold Fire provided a number of aircraft in support of CG 89, as well as support from an Allied Tactical Operations Center (ATOC).

On the ground, each side (Blue Northland and Red Southland) had one live division, with subordinate units operating in either FTX or CFX mode. One flank was simulated at brigade level using CBS, with scout platoons to provide the interface with the live portion of the exercise (see Figure 1). The other flank, the flank of the CBS simulation box, and the deep and rear battles for all units (both live and simulated) were simulated at the WPC.

The corps staff elements selected for training were primarily the intelligence staff and the deep battle or deep operations cell. This deep battle cell was set up to support Northland's operations. The WPC's GRWSIM model provided a Red threat from Southland to train the Northland deep battle cell. Note that in the simulations Southland used Red equipment, doctrine, and tactics, while in the FTX/CFX box, Southland's Gold forces used Blue equipment, doctrine, and tactics.

Caravan Guard 89: The Training Modes

During CG 89, some armored maneuver battalions participated in FTX mode, others in CFX mode, and still others in CAX mode.

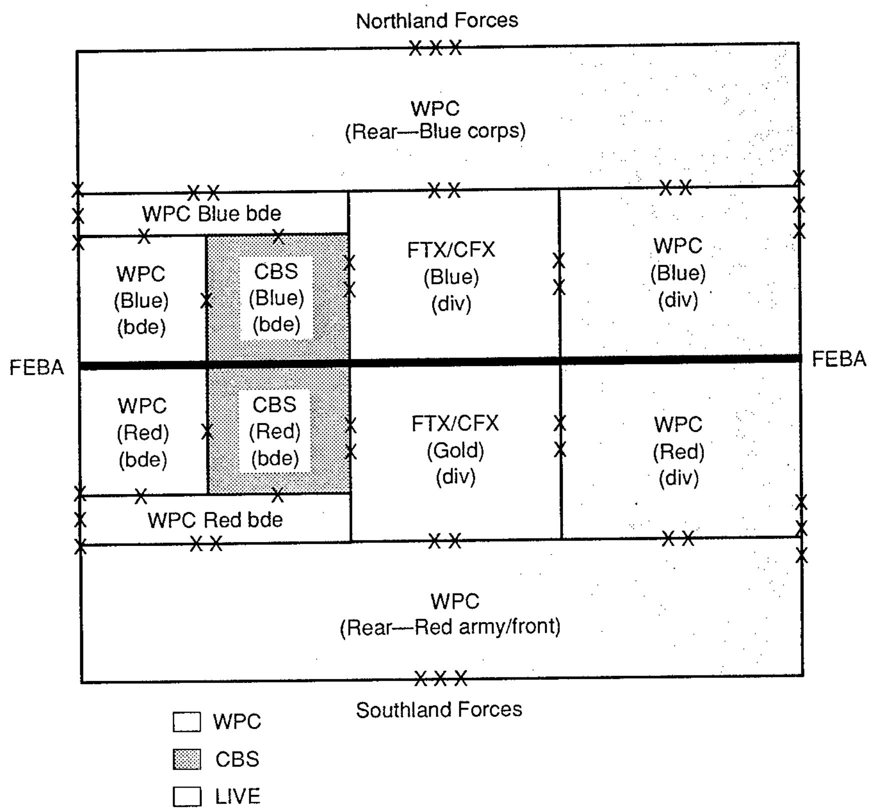

Figure 1—Exercise Battlefield for Caravan Guard 89

Umpires in the field had to assess not only the results of FTX units versus FTX units or CFX units versus CFX units, but also FTX units versus CFX units. This requirement made assessment somewhat difficult, since the umpires had to compare how one side *actually* employed its (FTX) forces to how the other side said it *would have* employed its (CFX) forces had they been present. In addition, the physical cues presented by all FTX vehicles being present tended to make that element easier to detect than its CFX counterparts with only command vehicles.

Furthermore, armored battalions could be employed in either FTX or CFX mode, but light infantry battalions needed to be employed in the FTX mode. Although maneuver damage was not a large problem for these infantry battalions, the assessment of results between FTX infantry and CFX armor was particularly difficult, especially at night, when the light infantry was most active.

Combining CFX and FTX training modes had some difficulties, but even more difficulties were faced when two types of CAX training modes were combined. Located on the flanks of the live-maneuver box, the simulations were included to create flank situations for the maneuver training audiences to address and to simulate enough units to flesh out a full corps, used to provide training for the corps headquarters elements. The accomplishment of both of these objectives was reasonably successful, indicating that a similar design would be useful for the upcoming CS 90 exercise. However, the interface between the two types of simulations, and between the simulations and the other training modes, did not work very well. Therefore, planning for CS 90 included ways to better interface the different types of training modes and thereby come closer to a "seamless battlefield." In a seamless battlefield, the seams between the different training modes are not obvious to the training audience.

Caravan Guard 89: The Sequence of Events

The sequence of events during CG 89 was planned as follows. During the first week, the Southland forces would attack northwards against the defending Northland forces. Then, during the transition weekend, the training audiences would reorganize their forces for the second week. During the second week, the Northland forces would counterattack southward against the defending Southland forces. This would allow both primary training audiences to be exercised in both offensive and defensive missions during the course of the two-week exercise.

Traditionally, the transition weekend of this exercise was performed in the administrative (noncombat) mode. In CG 89, it was decided that the live forces would be reorganized in the administrative mode but the simulated forces would continue to fight, so that a continuous "threat" would be presented to selected elements of the corps staff, such as the corps deep operations cell. The story behind the continuous transition weekend was that each of the live divisions was relieved by a simulated division late on Friday; late on Sunday, the live units would relieve the simulated units on-line.

Part of the objective of CG 89 was to exercise elements of the corps staff that were not usually exercised, such as the deep battle cell and the rear area operations center. Since CG 89 was a V Corps exercise, only Northland had these cells, and threats were presented to these cells during the course of the exercise. This design was employed in anticipation of CS 90, where there would be two opposing corps, each with deep and rear battle cells to be exercised.

In addition, there was an effort to play a flank threat in CG 89. Traditionally, live forces competed against each other in the live-play box, without threats from outside their areas of operation. In CG 89, it was decided that there would be simulated units along both flanks of the live-play box, so that each training audience had the opportunity to train in the tasks of coordinating with friendly flank forces and handling flank threats.

Centurion Shield 90: The Training Objectives

The training objectives for the battle staffs in CS 90 were as follows:

- Train the corps battle staffs to "look deep," plan and fight successful deep operations.
- Train corps battle staffs to effect coordination with flank corps in contact.
- Train corps rear area operations center (RAOC) to effectively plan and fight the rear battle on short notice.
- Train corps battle staffs to effectively manage close battle operations.
- Synchronize close, deep, and rear operations.

Centurion Shield 90: The Exercise Battlefield

Centurion Shield 90 took place in January 1990, drawing on lessons learned from CG 89. The geographical arrangement of the live and simulated training modes is presented in Figure 2. The live CFX/FTX "box" was in the southern portion of the full exercise box. The CBS box covered a brigade sector to the north of the CFX box. Actions to the north, west, and east of the live-play box were represented in the WPC models. South of the live-play box was assumed to be impassable terrain.

Centurion Shield 90: The Exercise Modes

Based upon the results of CG 89 and an unanticipated shortage in training funds, the plans for CS 90 were modified significantly. Rather than include a mix of FTX and CFX armored units, only CFX armored units were employed. This helped cut in half the estimated

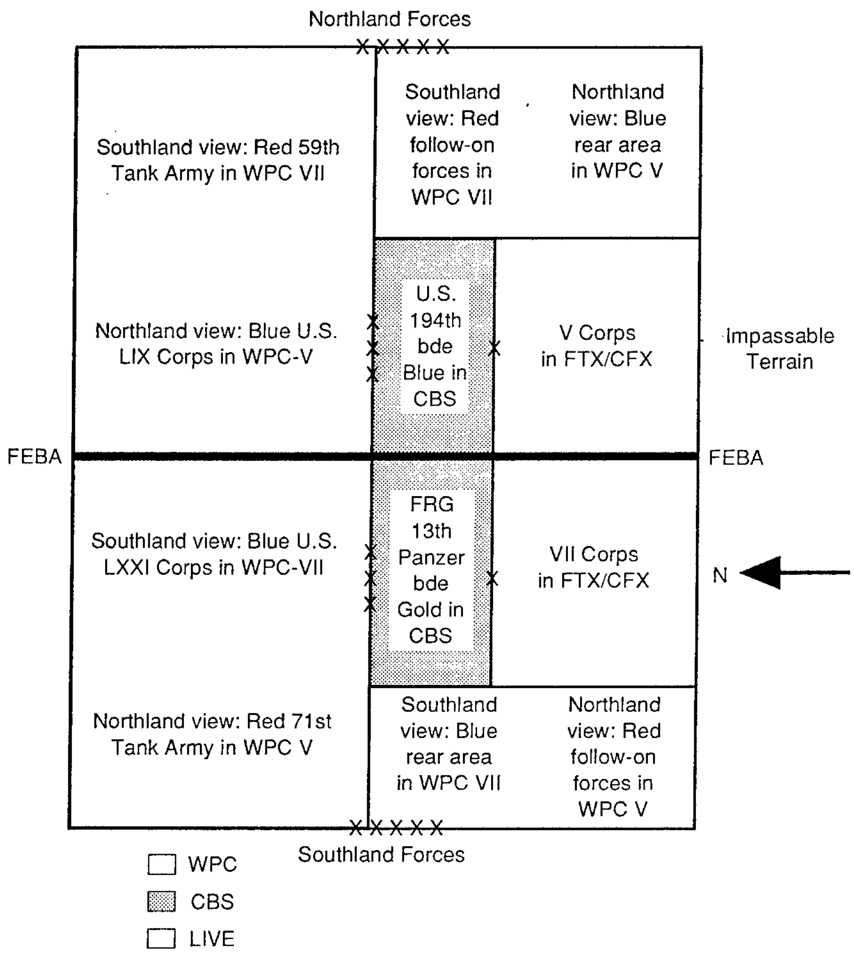

Figure 2—Exercise Battlefield for Centurion Shield 90

OPTEMPO costs (from $8.2 million for five M1 battalions in FTX, seven M1 battalions in CFX, three M3 battalions in FTX, and five M3 battalions in CFX to $4.31 million for all of these battalions in CFX). It also reduced the complexities associated with assessing movement and combat between FTX and CFX armored units. Light infantry remained in the FTX mode, giving rise to the same difficulties as in CG 89.

Simulations also played an increased role in CS 90, covering a wider area and including more simulated forces than in CG 89. The purpose was to take advantage of the potential opportunities to train functional areas not usually well trained in the more traditional modes. In addition, USAREUR's REFORGER Planning Group prepared a set of rules of engagement to define how simulated units from each model would interact with each other and with units in the live-maneuver box. (See Allen et al. (1992a), Appendix B, for further discussion.) The purpose of these rules was to present to the training audience as seamless a battlefield as possible. Although these rules of engagement did tend to make the battlefield seamless in some procedures, such as a call for fire, the overall assessment processes between training modes were not seamless.

The GRWSIM model covered a corps sector to the north of the CBS box, as well as the deep battle areas of the V and VII Corps. To present a Red threat to the corps deep battle staffs and flank liaison officers, two GRWSIM games operated simultaneously over the same geographical area. For example, the simulated corps (LXXI) north of U.S. VII Corps fought against a Red opponent (the 59th Tank Army). Meanwhile, the simulated corps (LIX) north of U.S. V Corps fought against a Red opponent (the 71st Tank Army).

Finally, two additional simulated corps were supposed to replace the "live" corps on-line during the transition weekend. Traditionally, the transition weekend is used to shift one live corps from the offense to the defense, while the opposing corps does the opposite. The transition weekend is usually performed in the administrative mode, so no combat occurs. However, since simulations were being employed, it was decided to have the simulated forces continue to fight through the transition weekend.[9]

Centurion Shield 90: The Sequence of Events

The sequence of events during CS 90 was planned as follows. During the first week, the Northland forces would attack westward against the defending Southland forces. Then would come the transition weekend, during which the training audiences would reorganize their forces for the second week. During the second week, the Southland forces would counterattack eastward against the defending Northland forces. This would allow both primary training audiences to be exercised in both offensive and defensive missions during the course of the two-week exercise.

---

[9]Since it is more difficult to relocate and refit live units on the ground than simulated units in the models, combat models allow the exercise to be performed continuously without an administrative break in the action.

The sequence of events was much more complex in CS 90 than in CG 89. Due to the increased scope of the exercise (two opposing corps in the training audience, as opposed to two divisions and selected corps cells), a number of experimental exercise design features were attempted. To present a Red threat to the deep battle, rear battle, and flanks of the training audiences, the concept of the overlapping Red and Blue games was employed, as described above. This degree of complication created considerable confusion for the training audiences, and it is not recommended for future exercises.

The exercise intended to present a continuous deep threat to each of the corps deep operation cells. Unfortunately, the sequence of committing deep forces became confused; a large number of these forces were committed early, and a continuous deep threat was present for only the first two or three days of the exercise.

There was also a continuous transition weekend in CS 90, similar in design and purpose to the transition weekend in CG 89. The story behind the CS 90 transition weekend was that the live corps on each side would be relieved in place late Friday evening by a simulated corps, which would in turn be relieved by the live corps late Sunday night. The purpose was to exercise the corps staffs throughout the weekend. Unfortunately, the method of implementation was to split the corps staffs, one to reorganize the withdrawn corps and the other to fight the simulated corps on-line. This was very confusing to the training audience. In addition, it is unrealistic to expect that two corps would relieve each other on-line in a 48-hour period. It might be possible for a division to be inserted and withdrawn in 48 hours if intensive combat is taking place, but it is not likely that a whole corps would be inserted and withdrawn again in 48 hours.

BENEFITS AND PROBLEM AREAS IDENTIFIED IN CARAVAN GUARD 89 AND CENTURION SHIELD 90

Both exercises were experiments as well as training events. Analyzing the results of the experiments revealed numerous benefits and problems both in the interfaces between the different training modes and in their representation of the various functional areas in the training audience. The overall assessment of CG 89 was that some of the modes appeared to have improved the training of some functional areas, but that the seams between the different training modes needed significant improvement. The evaluation of the CS 90 exercise concluded that the problems in the interfaces between training modes, combined with the increasing maneuver costs and restrictions, favored employing simulations as the primary training mode, with selected command elements in the field to achieve specific training objectives and for added realism.

Benefits of Caravan Guard 89

The training improvements appeared to occur primarily in the functional areas other than operations. For example, in the intelligence functional area, the use of simulations in a relatively freeplay mode provided better training for the intelligence staff. As opposed to a scripted exercise, in which the intelligence play does not have much effect, the intelligence play in the simulation made a significant contribution to the course of events. In particular, the deep operations cell faced the problem of monitoring enemy forces moving into the battle area and predicting their routes and arrival times. Using simulation to present the deep threat to the deep battle cell was an improvement over field exercises (FTX and CFX) that could not provide a deep threat because of manpower and cost constraints.

Of course, the different training modes posed other problems for intelligence staffs. As explained in more detail below, intelligence has difficulty training in the CFX mode, and to some degree in simulations, because of the simplified representation of units and simulated intelligence processes. Even so, simulations allowed significant improvements in the training of selected intelligence functional areas.

Another functional area that benefited was logistics. Detailed consumption rates, resupply requests, and the repair of damaged vehicles could be tracked well in the simulations. Simulated units could be constrained by limited ammunition and fuel or slowed at river crossings by a lack of bridging equipment. Conversely, FTX units tended to ignore simulated constraints on ammunition, and CFX units could ignore constraints on both ammunition and fuel. As a later section will describe, the tendency is to consider the time available to maneuver as too valuable to allow logistics considerations to constrain that maneuver.

One significant benefit to combining training modes was the event during CG 89 of live units withdrawing in the face of a flank exposed to simulated flank units. During the second half of the exercise, Southland was defending against the Blue counterattack. In the live-play box, it appeared to be holding. In the simulated-play box to its right flank, however, the defense was folding. Faced with an exposed flank, the Red commander called for a withdrawal of the live forces to the next defense line directly as a result of the simulated flank threat.

This was a significant event in that for the first time, events in a field exercise were being driven by the simulation. This was one reason why the flank play was designed to be so elaborate during CS 90.

After CG 89, it was hoped that the benefits observed there could be enhanced in CS 90, and that the problems encountered could be reduced. Overall, we concluded that the benefits of simulations might improve training in most functional areas above the battalion level. The difficult question was how to overcome the problems of each functional area with each training mode and with the interfaces between training modes.

Problems Encountered in Caravan Guard 89

Obtaining a Seamless Battlefield. The experimental nature of CG 89 encouraged the employment of multiple training modes. There were FTX forces, especially the light infantry units. There were CFX forces, which were primarily the heavy maneuver forces of armor, mechanized infantry, and armored cavalry units. And other elements were being trained in CPX and CAX training modes, primarily the commanders and staffs at brigade and above in the training audience. Also, live and simulated aircraft, both rotary- and fixed-wing, participated. One goal of the exercise was to try to make the presentation of the battlefield to the training audience as seamless as possible (i.e., the training audience should not be able to distinguish between live and simulated forces).

In CG 89, however, the seams were fairly obvious between the CFX and the FTX, between these two modes and the simulations, and between the two simulations. First, umpires had difficulty in assessing the outcomes of fielded units in the FTX and the CFX modes. A unit in the FTX mode has to maneuver all its vehicles on terrain subject to maneuver restrictions. Complications occur in facing, moving, and fighting every vehicle in a coherent manner. When both engaged units are in FTX mode, each is equally handicapped. But when one of the units is in CFX mode, this unit has a distinct advantage. Fewer complications and less friction occur when only a single vehicle is maneuvering rather than five units, or when five vehicles are maneuvering rather than twenty-five vehicles. Furthermore, the umpire must assess combat outcomes based upon how well one side actually arrayed its forces for battle versus a side that merely described how its forces would have been arrayed.

Second, there were problems in assessing interactions between live fielded units and simulated units. For example, a fielded unit has all the problems described above but even less information available on the location and strength of opposing units. A unit in CFX mode has about a fifth of the vehicles on the battlefield that the FTX unit has. A simulated unit has no vehicles on the battlefield. All the visual and aural cues of enemy presence are absent in the simulated mode. This makes it very difficult to assess combat outcomes realistically.

Third, problems arose in the interface between the two simulations. Because they lacked a sufficient interface, units from each of the simulations could not engage each other. Their presence could be noted as of the last update between the simulations, but there was no way to assess interactions among units in the two simulations. Updates typically occur every twenty minutes. Using extensive manual inputs, units in one simulation could be passed to the other simulation, but the different data structures led to translation problems in information passed between the simulations. As a result, there was a large and nearly impenetrable seam between the two simulations, as large a one as between any of the other training modes.

Functional Area Difficulties. The imperfect interfaces were not the only difficulties experienced with exercise design. Difficulties also occurred within the functional areas, primarily in the staff functions of intelligence, rear operations play, logistics, and personnel.

Intelligence. A number of problems vexed the intelligence play of the exercise. Live intelligence assets had more difficulty detecting CFX units than FTX units. Simulated intelligence assets had difficulty detecting live units due to the delay in updating live unit locations in the simulations. Live units close enough to be in contact with simulated units could not report on simulated units, or vice versa. A deep threat was played in the simulation, but the movement rates allowed caused the follow-on echelons to move toward the FLOT too rapidly to allow the normal intelligence cycle to operate effectively. By the time the various taskings went out, the threat had already moved beyond their search areas.

Rear Area. The insertion of threat air assault forces into the rear area of the training audience occurred as a scripted event during CG 89. This created a threat to the rear area, thereby forcing the appropriate portions of the training audience to react. Due to the artificialities of scripting the insertion of these forces, many elements of the training audience were disturbed by the lack of realism. The intelligence staffs were given no warning, the air defense units did not have the opportunity to engage the forces on ingress or egress, and the RAOC was given no forewarning of the insertion.

Logistics. Nor was the logistic area free of difficulties. For example, maneuver and artillery forces frequently face a shortage of simulated ammunition. Although the effects of the combat service support (CSS) functional areas were understood to be important, constraints from the CSS functional areas were not allowed to interfere with maneuver and fire support functional area training. As a result, the combat arms continued to train as though they were unconstrained by the realities of ammunition shortages or other CSS constraints.

The reason usually cited for this is that maneuver time is considered too valuable to waste on waiting for realistic CSS. The OPTEMPO costs are high when maneuver forces are actually in the field. It takes a lot of time to plan and coordinate the use of maneuver areas, whether they are battalion-sized training facilities or the (more or less) open countryside used during REFORGER exercises. As a result, most exercise objectives are for the G3 and then the G2, usually at the expense of the training objectives of the G1 and the G4. Logistics and personnel constraints are not allowed to significantly slow or stop maneuver training, despite repeated efforts to represent the effects of CSS over many exercises. During both exercises, the planned influences of CSS were overridden by the maneuver and fire support elements. As long as forces in the field are present in an exercise, the tendency will be for the G3/G2 tasks to dominate and to override the effects of G1/G4 tasks.

Benefits of Centurion Shield 90

CS 90 benefited from the experience of CG 89 by learning where the problems were in the training mode interfaces. A number of steps were taken to smooth the seams between the modes and simulations and to enhance the functional play. For CS 90, the REFORGER Planning Group came up with a set of rules of engagement for how forces in one training mode could communicate with those in another, thereby attempting to create a seamless battlefield. Unfortunately, these rules of engagement were only partially successful, leading to a battlefield that was still primarily seamed between the training modes.

The main benefit resulting from CS 90 was that this second experimental exercise helped determine which combinations of training modes work well for higher-echelon (division and above) exercises and which do not. Overall, we concluded that field exercises are currently best for training at battalion and below (including a brigade headquarters and a brigade "slice" of support). For training above battalion, CAX is probably the best training mode for simultaneously training all major functional areas. Mixing field exercise components (FTX and CFX) with simulations causes too many interface problems to be a worthwhile method of meeting the training objectives. Therefore, the lowest echelon that should be included in the training audience in a large-scale multiechelon exercise is the battalion headquarters. For example, a single exercise may include corps, division, brigade, and battalion staffs, but not company or below. The highest echelon that should be included in regular field exercises with maneuver elements in the field should be the brigade.

This is not a rule etched in stone. There may be reasons to exercise a unit larger than a battalion or brigade in the FTX mode. However, one should understand that the main emphasis of an FTX is on the maneuver elements and the G3 staff element, rather than on all of the functional elements and their interactions.

The functional areas that benefited in CG 89 also benefited in CS 90. But some of the exercise execution problems mitigated some of that benefit. For example, owing to exercise sequence difficulties, the deep operations cell was not presented a continuous threat. This was not a limitation of the training mode but a problem in exercise design and execution. Overall, compared to traditional FTXs, simulations provide an opportunity to better train different functional areas at the same time. The reasons for this are explained in more detail below.

Problems Identified During Centurion Shield 90

There were three main problem areas encountered during CS 90: the interfaces between the training modes caused seams in the battlefield; the training of some functional areas was hindered by training mode artificialities; and the need for the maneuver elements to maximize the use of the time in the field dominated the training objectives of some of the other functional areas.

Battlefield Still Not Seamless. The results of CS 90 showed that the procedural issues, such as calls for fire and other support requests, worked fairly well under the "rules of engagement" process. However, the assessment processes across different training modes did not work as well. More importantly, the seams were apparent to the training audience during the exercise, especially the seams between field units (FTX and CFX) and simulated units (CAX). Also, CS 90 experienced difficulty in some of the functional areas. For example, even though the intent was to have forces face each other in the same mode as shown in Figures 1 and 2, the tendency is for flanking operations, so that forces from one mode will often try to encounter forces from another mode. Unless their interactions are well defined, a combination of problems will arise, resulting in unrealistic interactions between forces represented by different training modes, as described below.

Difficulties always occur when field units play the exercises in different modes. The FTX units must interact with CFX units, but they cannot do so on an equal basis. During CS 90, light infantry units in FTX mode infiltrated CFX mechanized units at night. The physical lack of vehicles and personnel in CFX units allowed light infantry to infiltrate easily.

Interactions of live units with simulated units posed even greater problems. FTX units knew simulated units could not affect them in direct combat, so they ignored them. In addition, simulated units need to interact with both CFX and FTX units. But how do they interact in close combat? There is no representation of simulated units on the ground, so how are they detected or their effects assessed? Similarly, live aircraft had to fly against live targets (to meet Air Force training objectives), so certain types of aircraft could not fly against certain types of targets.

Furthermore, simulated units cannot be detected by live sensors. Does one then represent all intelligence collection via simulation, including the live units? If so, how does one ensure the timely update of unit locations into the simulations? During CS 90, reported unit locations were often two to four hours behind actual unit locations, even to the exercise control staff.

During CS 90, CFX/FTX unit locations were reported by the umpires to the Umpire Control Center (UCC), which sent the data to the Exercise Control Center (ECC). At the ECC, the CFX/FTX unit locations were entered manually into CBS and then passed to GRWSIM through the simulations interface. Once units were in the WPC model data base, the Intelligence Collection Model (ICM) was run on that data base to detect unit locations, including the CFX/FTX units. Due to the delays associated with each step, the intelligence data on CFX/FTX units were at least two to four hours old by the time the corps G2 received any of it. However, the data collected on the simulated units were more continuous, since the computer updated unit locations about every 20 minutes. As a result, the intelligence cells observed that location data collected on the CFX units were quite "jerky," whereas the location data collected on the simulated units were more continuous. It was thus unusually difficult for the G2 staffs to track the live units through the simulations.

In virtual contradiction to this situation, the CFX/FTX players came to rely more on the ICM reports passed to them from corps than on the live reports from actual sensors in the field. This is because the ICM reports were more accurate than the reports from the live sensors, in spite of the time delay. This apparent paradox can be explained when one considers that the ICM reports were on battalion-sized units, while the field reports were—essentially—on individual command vehicles posing as companies or batteries. The aggregation process in the simulations reduced a lot of random error on the individual vehicle locations; hence the center of mass reported by ICM was, even though late, more accurate than the individual field data points available from CFX/FTX units.

The interface between the two simulations did not work well either. For example, a unit from the WPC model engaged by a unit in the CBS model would always be destroyed by the latter, regardless of the size and type of units engaged. Fire missions passed between simulations would have to be reinserted by hand. Air defense assets were duplicated in both models since neither model's aircraft were affected by the other model's air defenses. When air defense assets were suppressed or destroyed in either model, this effect had to be inserted by hand in the other model.

Another exercise design problem was that the simulated corps to the north was too elaborate for the benefits derived from creating a potential flank threat. Because of the scope of the exercise, the two opposing corps headquarters in the training audience focused their attention on their own units rather than on the flank corps. Since the flank corps was not under their command (as was the case in CG 89, where the simulated flank division and the live division were under the same corps headquarters), it received little attention. In addition, the seams were apparent to all in the training audience. For example, live aircraft could not fly in the simulated box, and simulated aircraft could not fly in the live box. Since simulated ground forces could have little effect on the live battle, simulated units were, for the most part, ignored by the live headquarters.

As a final example, CS 90 experimented with providing a simulated Red threat against elements of each Blue training audience, even though live Blue training audiences were opposing each other in the same exercise. One part of the battlefield involved opposing live Blue elements of the training audience, while other parts of the battlefield involved simulated Blue versus Red threat elements. To make matters worse, there were two simulated games superimposed on each other, so that each live Blue corps had a simulated Blue corps on its flank that faced a Red threat. The extensive confusion that followed from this arrangement was not unexpected, and we recommended that this technique not be employed in future exercises.

Overall, the battlefield was not seamless, nor could anyone determine a simple way to make it so without "instrumenting" every vehicle in the field (such as at National Training Center or the Combat Maneuver Training Center). Thus a number of artificialities resulted, thereby reducing the realism and the training benefit of the exercise.

Functional Area Problems Persist. During CS 90, there remained problems in many functional areas. Major problems occurred in the following three areas: intelligence, deep operations cell, and the dominance of G3 over other functional areas.

Intelligence. An important problem in a CFX is that the play of live intelligence becomes significantly limited. Live sensor platforms can detect only live units in the field, FTX being easier to detect than CFX. It is much more difficult to find a single wheeled vehicle representing a tank battalion in CFX mode than it is to find an actual tank battalion. In addition, when a scout does detect a single wheeled vehicle, is he detecting a tank battalion or a mechanized battalion? What is the strength and posture of the "force" just detected?

There were similar problems with reporting detections between live and simulated units close enough to detect each other. This was a problem in CG 89 and remained a problem in CS 90. However, updating the locations of live units in the simulations was somewhat improved in CS 90, as a result of stronger emphasis on getting those reports in on a regular basis. There were still problems in determining actual versus reported location, as is usually the case on any battlefield. However, the simulations must rely on the reported locations from the live units, since there was no objective way of determining whether or not a false location was being reported accidentally.

Deep Operations Cell. The deep operations scenario started out well, and for most of the first week the deep operations cells were presented an adequate threat. They were defining named areas of interest (NAIs) and target areas of interest (TAIs) according to doctrine, and they were generally tracking the threat units. However, after the initial (simulated) enemy reserves were committed, the deep threat disappeared and the deep operations cell had nothing to monitor. Red forces were either committed to close operations, or they were out of the corps area of influence or interest. For example, since live aircraft could not fly against simulated targets, when enemy forces appeared in the rear of the flank simulated corps, the targets had to be passed to the adjacent simulated corps.

G3 and Maneuver Elements Still Dominate Other Functional Areas. Large-scale multiechelon field exercises have a tendency to ignore the effects of some functional areas (e.g., G4 and G1) in order to train another functional area (e.g., G3, maneuver elements, and G2). As mentioned above, the value of maneuver time is usually cited as the reason. This is not a trivial concern. Opportunities to move large units over real terrain, even under maneuver restrictions, are comparatively rare and expensive. To keep the maneuver troops waiting for simulated supplies is more "realistic" training, but it is not necessarily cost-effective. As a result, maneuver element and G3 concerns have tended to dominate the training needs of the other functional areas.

This does not mean that the other functional areas are not important or that they receive little or no training. On the contrary, most combat support and combat service support elements do perform their own field exercises. The problem is that if the training objectives of the other functional areas conflict with the training areas of the maneuver elements and the G3 staff, the latter wins the conflict.

From these priorities and our observations of CG 89 and CS 90, we have concluded that the FTX and CFX training modes will, by their nature, continue to be dominated by these priorities. As long as maneuver elements are in the field, the concern of maneuver time being too valuable to constrain for logistics considerations will continue.

Conversely, computer-assisted exercises do not suffer from the constraint of having live forces in the field. There is little "opportunity cost" involved in keeping simulated maneuver units waiting for supplies. The realism of this type of situation will add to the training experience for all functional areas—G1 through G4, rear and close operations, ground and air coordination.

As a result, we recommend that for large-scale multiechelon exercises, maneuver elements of the battalion or below should not be included in the training audience. Battalion staffs are the lowest level in the training audience, and they may actually serve best as response cells. (See the later section on the balance between training audience and training support personnel.) This does not mean that the battalion staff should not be in the field. On the contrary, the ability to move the battalion headquarters and tactical operations center while maintaining continuous command and control is an important training objective. The difference is that without the maneuver elements in the field, there is no urgency to use exercise time for maneuvering at the expense of the other functional areas.

We expect that simulations will provide the best vehicle to simultaneously train all functional areas in multiple echelons above the battalion level. In the next section, we elaborate further on this subject and discuss some of the planning and tradeoffs involved in designing this type of exercise.

3. EXERCISE DESIGN OBSERVATIONS RESULTING FROM AN ANALYSIS OF CARAVAN GUARD 89 AND CENTURION SHIELD 90

Analysis of the two exercises leads to four primary suggestions for future exercise planning.[1] First, the commander and the designers of the exercise should establish and clearly articulate in sufficient detail the training objectives for the exercise. Second, they should, whenever possible, select the single exercise mode that provides the greatest opportunity to achieve all of those objectives. Third, the design of the exercise should be evaluated to ensure that it allows for all the training objectives of the many functional areas to be met, and that it sufficiently engages the training audience in each of those functional areas. Finally, the exercise designers must decide what balance they are going to strike between freeplay and scripted exercise, and what tradeoffs must occur between the training audience and training support personnel.

SELECTION AND ARTICULATION OF TRAINING OBJECTIVES

The first step in exercise design is to select the training objectives. The nature of exercise design makes this an iterative process. The achievement of some training objectives may not be feasible given the constraints, such as time, personnel requirements, location, and training mode artificialities.[2] Since the elements of the iterative process are unique to each exercise, we will not go into detail here, but exercise designers should keep them in mind. The first subsection will focus on the need to define training objectives specific enough so that the exercise can be designed and executed to ensure that they are met.

Just as important as defining the training objectives is the need to articulate them clearly. Each of the different training elements, including every functional area in each echelon, must know what is ex-

[1]For specific suggestions for better representing each functional area and the interfaces between training modes, see Allen et al. (1992a, 1992b).

[2]For example, a live-fire exercise requires the availability of adequate supplies of ammunition, as well as a prepared live-fire range. The scheduling of units through limited training facilities creates the need for training schedules with very long lead times, sometimes measured in years.

pected of it. In addition, each training element must know the priority of its training objectives with respect to other training elements. This is important because in large-scale multiechelon exercises, one element's training objective may compete with another's. Satisfying both training objectives may define an impossible problem, given the constraints. Therefore, at some point, selected elements of the training audience may need to become training support elements to achieve a higher-priority training objective for other elements.

The Need for Specific Training Objectives

The commander plans his training program around his unit's mission with a battle focus. Battle focus defines a peacetime training program based on likely wartime missions, understanding that no unit will be able to be fully proficient in every task all the time. Based on this approach, the Mission Essential Task List (METL) prioritizes the tasks the force commander considers most important to train, subject to the unit's likely wartime missions, its current capabilities and deficiencies, and its personnel rotation schedule. The METL subsequently drives the selection of the types of exercises desired, the balance between the various tradeoffs, and the modes of training to be employed.

The success of employing the METL has apparently been very good at division and below, and especially in battalion-level training. However, difficulties have been observed in defining the METL adequately at corps and above.

The reason for these difficulties appears to relate to both the broad scope and the complexity of interrelated tasks inherent in higher echelons. For example, the number of tasks essential to a battalion-sized unit is fairly well defined. Missions such as "hasty defense" or "river crossing operation" can be described in relation to the assets organic to the battalion and the assets or units "chopped" to the battalion in support of such a mission. The METL at battalion level tends to be focused on the following:

- U.S. assets and organizations (rather than multinational).
- The basic staff functions associated with a small staff.
- Noncombat functions associated with sustaining a small force consisting of a few types of assets.
- A time horizon associated with battalion operations.
- A set of likely missions that tend to be maneuver-oriented.

In contrast, corps and above echelons tend to have a larger number of tasks that could be considered for the METL. The METL at corps and above could be focused on the following:

- Multinational assets and organizations, rather than U.S. only.
- The staff functions associated with large staffs, with many specialized elements.
- Noncombat functions associated with sustaining a large force with many types of assets at various stages of availability.
- A time horizon associated with both current and future operations.
- A set of likely missions that tend to be more sustainment and less maneuver-oriented.

For example, during CS 90, a simulated U.S. division was being assigned to an on-line U.S. corps. The division thought it was a U.S. asset before being assigned to the corps, while the Army Group considered the division to be a NATO asset. As a result, confusion arose regarding which set of procedures should be followed: assigning a NATO asset to a U.S. corps or assigning a U.S. asset to a U.S. corps. The complexities associated with handling multinational forces do not usually arise at the battalion level.

Corps and higher staffs tend to include a number of specialized staff functions, such as the deep operations cell in a corps staff. The job of this cell is to watch for enemy forces behind enemy lines but within the corps' area of interest and to plan and coordinate efforts to delay and disrupt such forces. Similarly, rear area operations centers (RAOCs) monitor friendly territory for enemy incursions and plan and coordinate efforts to disrupt and destroy such forces. Battalions, owing to their size and mission, do not include specialized staffs.

Sustainability issues tend to gain more attention at the higher echelons than at the lower. The higher the echelon, the more attention is focused on providing adequate equipment, supplies, and personnel for sustained operations. Higher echelons have many types of assets and personnel to sustain, not just a few, such as at battalion level. In addition, staff members at higher echelons are more concerned with the return of equipment and personnel over longer planning horizons.

Planning horizons vary by echelon. At lower echelons, the planning horizon is measured in hours, up to 48 to 72 hours at division level. At higher echelons, planning horizons extend from days to weeks, depending upon the echelon. Because of the longer planning horizons at higher echelons, their plans tend to be more complex. The coordination between component staff elements, including the specialized staffs and liaison officers, must occur across longer time frames, over wider geographical areas, and among more abstract issues such as air space control coordination.

All of the preceding examples become even more complicated in multiechelon exercises. Not only are there more specialized staff elements at each echelon, but also each of these specialized elements must coordinate with higher, lower, and adjacent elements in other headquarters. It should not be surprising, then, with all of the real-world complexity associated with higher-echelon (corps and above) missions, that the METL for higher echelons is frequently not defined adequately when preparing for higher-echelon exercises.

For example, during CS 90, one component training objective was to "perform an air assault operation." This objective was not sufficiently specific, since it identified neither the echelon responsible for performing the air assault operation nor the specific functional areas to be trained. It did not define the desired depth of penetration, the mission of the ground element once inserted, the duration of the mission, and how it would affect sustainment issues.

Not only does the principle of METL apply at higher echelons, *it is even more important to define the METL specifically for the higher echelons.* The METL is a priority list of what is important to train. Since both the scope and complexity of the possible missions are greater at the higher echelons, it is therefore most important to clearly prioritize that list at the higher echelons.

The problem, then, is how to encourage and enforce the adequate definition of the METL at the higher echelons in developing their training programs. Traditionally this has not been accomplished for the higher echelons and functional areas in multiechelon exercises. In the absence of a full METL, an alternative approach might be to develop a METL for each training element that has less detail than a fully developed METL but more detail than the broad training objectives of CG 89 and CS 90. Completing a METL assists in the planning, execution, and evaluation of an exercise.

Clearly Articulating Specific and Measurable Training Objectives

Once the training objectives for a specific training event have been defined, they must be clearly articulated. Lack of sufficient detail in the training objectives affects the exercise design, execution, and evaluation. The result of poorly articulated training objectives is that some of the objectives may not be accomplished. For example, the five main exercise objectives for CS 90 were to:

- Train the corps battle staffs to "look deep," plan and fight successful deep operations.
- Train corps battle staffs to effect coordination with flank corps in contact.
- Train corps RAOC to effectively plan and fight the rear battle on short notice.
- Train corps battle staffs to effectively manage close battle operations.
- Synchronize close, deep, and rear operations.

Although some may consider these objectives specific, they are, in fact, very far from the detail necessary to ensure that they could be met. In fact, not all of them were met in CS 90.

In the case of the deep battle, for example, the presentation of the threat to the deep operations cells of each corps was not designed to be continuous throughout the course of the exercise. One or two enemy divisions did appear in the corps area of interest during the first two days of the war. But from the third through tenth days, no additional enemy forces appeared in each corps' area of interest. As a result, there was no "deep battle" to synchronize with the close or rear battles, nor did the deep operations cells receive adequate training for most of the exercise.

Similarly, the rear area operations centers were not provided an adequate threat, and no rear battles materialized. This happened mainly because there was no incentive for either corps to attack targets in the enemy's rear areas. Since CSS did not constrain the movement or combat rates of maneuver forces, there was no combat benefit from attacking the enemy's rear. As a result, the synchronization of the close, deep, and rear operations could not be trained, since there were no rear or deep operations to coordinate.

We suggest that the definition of adequately specific training objectives would help ensure that they are met. More specific training objectives will influence the planning, execution, and evaluation of the exercise.

During the exercise planning phase, a unit's METL, full or reduced, will guide the planner in the types and sequence of situations that should occur to meet the training objectives. To improve on the preceding example, a schedule of the minimum amount of force over time to be presented to the deep battle cell could be defined. If the freeplay aspects of the close operations preclude the insertion of another division, then possibly other forces could be inserted to challenge the deep operations cell. For example, an independent tank regiment, an MVD brigade, a long-range artillery battalion, or a major convoy could be entered into the corps' area of interest to sustain the presentation of the threat to the deep operations cells. These additional elements might not even be heading for the close battle area, but might simply be moving laterally across the corps' area of interest.

During the exercise execution phase, the exercise controllers will be monitoring the achievement of specific training objectives. To continue our example: if no threat has been presented to the deep operations cell for two days, and the opposing side has no intention of entering additional forces as part of his freeplay maneuvers, then the exercise controllers may opt to enter one of the alternative forces listed above. These actions will not disrupt the freeplay aspects of the close battle, but they will maintain the presentation of the threat to the deep operations cells.

Not every possible task needs to be trained during an exercise. The METL could be grouped into categories where at least four of the five most important tasks are achieved, and seven of the ten secondary tasks will be achieved. This provides some flexibility to the control staff to accomplish some alternative training objectives and not others, depending upon the way the freeplay part of the exercise flows. (See the section below on balancing the tradeoffs for more discussion of the issue of freeplay versus scripted exercise components.)

Let us stress this point: *It is essential that the training objectives of each training element be prioritized.* A training element is one functional area in one echelon. In large-scale multiechelon exercises, there will be many training elements, each with related but slightly different training objectives. As described in the next subsection, each training mode will entail artificialities. Any one artificiality may be compensated for, but a compensation may create additional artificialities in other functional areas. Satisfying every training element's training objectives at the same time may be an impossible task given the constraints of this complex situation. Prioritizing the training objectives across all training elements helps prepare both the exercise directors and the training audience for how these compensations will be implemented.

For example, if a simulated unit needs to be moved from one part of the battlefield to another by a "magic move" for purposes of accomplishing a high-priority training objective, then the unit may be moved in the computer. At the same time, however, this event presents an unrealistic picture to the intelligence training element at one or more echelons. If a suitable explanation cannot be provided to the affected intelligence cell, such as a plausible deception operation, then the intelligence cell(s) will need to be drawn temporarily into the training support function for purposes of supporting a higher-priority training objective.

The benefits of such a procedure go a long way in presenting a realistic and beneficial training experience. Currently, when a magic move occurs, the intelligence element finds out about it when either their sensors pick up the change or the maneuver elements complain that they have just been surprised by an enemy force that came out of nowhere. Both the maneuver elements and the intelligence staffs suffer from such a direct intervention. By cooperating with the intelligence element, the directing staff and the intelligence cell can provide the maneuver element with a plausible story, at the same time ensuring that the intelligence staff does not get surprised by an unrealistic event.

SELECTING THE NUMBER AND TYPE OF TRAINING MODES

Each training mode has advantages and disadvantages with respect to the costs and the training benefit provided to each functional area and echelon. The real issue, however, is not which mode is best for which functional area or echelon, but rather *which mode best achieves the training objectives for the combination of functional areas to be trained in each echelon.*

One can train any single functional area in one echelon using any training mode, because one can compensate for the artificialities of any one mode by controlling the information flow to the primary training audience. But when many functional areas are trained, especially across many echelons, the compensation techniques tend to interfere with each other.

The following case illustrates a single training mode used to train a single functional area in a single echelon. The Ordnance School uses a manual "wargame" CPX to train officers during their branch basic course at Aberdeen Proving Grounds. The controllers provide messages to the commander and staff of a maintenance unit, including the maintenance workload, priorities from higher commands and supported units, the degree of threat, and the resources available.

The training audience must allocate resources to meet the demands, constrained by resource supply levels and perceived degrees of acceptable risk. Unexpected events are inserted during the course of the exercise. All information flows into and out of the training audience are handled by the exercise controllers. This type of exercise works well, since there are no conflicting training objectives nor artificialities imposed from other functional areas, echelons, or training modes.

But even under a single training mode, artificialities can occur between functional areas. For example, a CFX is often employed because it is less expensive than an FTX. Only one vehicle represents a platoon or even a company. However, the intelligence collection functional area suffers from the lack of a realistic number of targets to detect in the field. The lack of real vehicles makes it difficult for live sensors to detect them. The situation is even more difficult when wheeled vehicles substitute for tracked vehicles, as they did in CS 90.

For another related example, forces in the field in either the FTX or CFX training modes are allowed to "resurrect" approximately four hours after becoming casualties. Therefore, every four hours, units that are wiped out or severely attrited reappear as if by magic. The intelligence collection and fusion functional areas are again handicapped by this artificiality. Intelligence cells attempt to monitor the status of enemy units, as well as enemy efforts to reconstitute lower-status units. The automatic reappearance of enemy forces with no evidence of reconstitution and no continuous-status picture over time causes an unrealistic constraint to be imposed upon the training of intelligence collection and analysis cells.

Note that the preceding examples illustrate the effects of a single training mode on different functional areas. Additional problems arise when multiple training modes are applied in a single large-scale exercise, as presented earlier in Section 2. The more training modes used in an exercise, the more interface problems arise. The number of problems increases geometrically, since each functional area must interface with every other functional area across multiple interfaces. The combination of interface problems and the need to compensate for all the inherent artificialities make the use of multiple training modes in a single exercise unattractive.

In Section 2 we concluded that a large-scale multiechelon exercise should not include maneuver elements in the field, but may involve moving battalion headquarters and tactical operations centers in the field. The primary training mode will be a single simulation in order to keep the interface problems to a minimum.

Whenever two or more training modes are employed in a single exercise, the interface problems become very apparent. The more training modes employed in a single exercise, the more interface problems can occur. As a result, we conclude that it is not worth the effort to try to handle or compensate for all of the interface problems between field and simulated components. Partly for this reason, USAREUR currently plans that REFORGER 92 will be primarily CAX, with units represented down to battalion (and possibly company in some cases) headquarters.

CRITERIA FOR PRETESTING AN EXERCISE DESIGN

We define pretesting to be the actions taken to "step through" the exercise during the exercise design process to ensure that the training objectives are likely to be met. There are two main components to this pretest: pretesting the exercise from the *viewpoint* of each functional area training element to ensure that its focus of attention will be on the training objectives; and pretesting the likely *sequences of events* to ensure that each element of the training audience has the opportunity to meet its training objectives. Although both of these components can be examined during the same pretest, the first tends to be time independent, focusing on the organization of the training audience and the training modes. The second component includes the effects of time, and it focuses on the sequence of events perceived by the training audience as presented by the training modes over time.

Both of these procedures involve examining the exercise from the perspective of the different elements in the training audience, rather than from the perspective of the exercise designer. The purpose of this pretesting is to identify and eliminate potential problems where the training objectives of each training element may not be met because of a faulty sequence of events or because the training audience's attention has not been engaged.

Exercise Design Must Demand the Attention of Each Functional Area

Pretesting the exercise design from the perspective of each functional area is desirable to ensure that each training element's focus of attention will be on the training objectives and not on the artificialities created by the training mode(s) selected.

For example, the design of CS 90 failed to require adequate attention to events on the flank. The flank play in CG 89 was more successful, since the simulated flank units were part of the same corps as the live units. The corps commander paid attention to both the live and the simulated battles in CG 89. Unfortunately, the corps headquarters in CS 90 consisted of all live units except for an attached brigade to the north in the CBS box. Due to the interface problems between the live and the simulated boxes, little interaction took place between the live and simulated play. (The notable exception was the TACFIRE requests for fire interfaced directly with the CBS computer model.)

As a result, there was nothing to force the corps commanders to pay any significant attention to the corps to their north. The manpower requirement to run a whole simulated corps (including the opposing side and two shifts for 24-hours-a-day operation) was about 200 augmentee personnel; along with the expensive communications link, the flank simulated corps was not worth the cost or the effort in CS 90. It would probably have worked better to have a simulated division as part of the corps training audience, to ensure that the training audience was forced to pay attention to the flank play.

A similar example occurred with presenting the threat to the deep and rear operations cells, as described below. From the viewpoint of the deep and rear operations cells, the sequence of threat presentation is the key factor to gaining and maintaining their focus of attention. The exercise's failure to command attention in both of these areas meant that three of the training objectives were not fully accomplished.

Overall, live forces did not feel that they had to pay attention to the simulated forces in CS 90. Except in a few cases, live forces could be affected effectively only by live ground and air forces. Similarly, simulated forces perceived that for the most part, they had little effect on live units, and therefore focused their attention on simulated units. Even between simulations, interface problems caused units from different simulations to avoid each other. As a result, problems with the interfaces between the training modes allowed many functional areas to ignore other functional areas.

Pretesting the exercise design from the viewpoint of each functional area element does not require that a full mock-up exercise be executed, nor that all of the functional areas be pretested at the same time. Some pretesting is already accomplished, but this is usually performed by exercise design personnel who understand how the whole exercise is supposed to work; they do not do it from the perspective of any single functional area element. The exercise designers have too much information and may be too close to the problem to identify how a player will naturally act in a given functional area element.

To adequately pretest the exercise design, a person or persons not involved with the overall exercise design but familiar with that functional area should be placed in the position of the player for the forthcoming exercise. The exercise designer can present to the test-player the scenario that will be played, along with a description of the information presented, the form and frequency it will be presented in, and any known artificialities in that presentation. For example, an intelligence cell may not be receiving raw intelligence data, but data that have already been processed by another cell and not presented in the form of a familiar intelligence overlay.

The test-player should be questioned as to ways to compensate for these artificialities, which tasks may be trained, and which tasks should be considered training support of other functional area elements. For example, if the information coming in provides no uncertainty, the intelligence cell may be asked to prepare an intelligence overlay that presents a more realistic level of uncertainty.

Most importantly, the test-player needs to determine whether or not the scenario as defined will force his attention on the training objectives. For example, if the player knows that simulated units cannot interact frequently with live units, he may not be able to coordinate activities between live and simulated units except in rare cases. Therefore, a simulated flank corps not under the command of the test-player will be identified as unable to force the interest of the test-player on the flanks.[3]

As part of the pretesting from the viewpoint of each functional area element, the test-players should identify the frequency with which the information must be provided in order to accomplish their tasks. For example, the frequency of presentation of the threat to deep and rear operations cells and the flank players is essential to accomplishing their training objectives, as described further below. Similarly, the rate at which the training mode updates information may not be adequate to achieve some of the tasks in certain functional areas. For example, monitoring the threat from the deep operations cells requires an update frequency that allows virtually continuous monitoring of the named and target areas of interest. Otherwise, the gaps in

[3]During the planning of CS 90, the intended interface under development was intended to be more robust than it turned out to be. The magnitude of the difficulties in interfacing the training modes and the effort to provide those interfaces on a "by exception basis" were not fully understood until after the experimental exercise was executed. See Allen et al. (1992a), Appendix B, for a more detailed discussion of which interfaces worked and which did not during CS 90.

the data flow may provide an unacceptable artificiality in accomplishing the tasks.

Pretest Sequence of Events To Ensure Training Objectives Are Met

Pretesting the sequence of events involves "playing through" the exercise in its likely sequences of events to determine from the perspective of each functional area whether the training objectives are likely to be met. Unlike pretesting from the viewpoint of each functional area separately, there is a strong advantage to pretesting the sequence of events simultaneously from most or all of the functional area elements.

For example, one of the objectives of both CG 89 and CS 90 was to train the deep operations cells. During CG 89, the deep threat moved so quickly through the NAI and TAI that the deep battle cell had no opportunity to fulfill its function. Similarly, the RAOC was unable to respond to the enemy air assault in the rear because the sequence of events was not detailed enough to allow it to react.

During CS 90, the deep operations cell was presented a threat in its area of interest for only the first two days of the exercise. The sequence of events did not ensure that a continuous threat would be presented to the deep operations cell. Therefore, the deep operations cell received little training for the third through tenth days of the exercise. Similarly, the presentation of the threat to the rear operations cell was also limited. The size of the threat force that survived the insertion was too small to be militarily significant.

As a result, the structure of the sequence of events was such that the training did not occur. One of the reasons is that the simulated units in the rear of each simulated Army-level response cell were needed by other maneuver elements earlier than expected. Not realizing what effect the early release of these units would have on the training of the deep operations cells, the Army-level response cell released the units to subordinate commanders. As a result, no simulated forces were left to present the threat to the deep operations cell. Had this aspect been pretested simultaneously with other functional area elements under "what if" conditions, then the lack of a continuous threat presented to the deep operations cell may have been detected.

Of course, not all of the problems that could arise in an exercise can be identified through this method. It is always easy to say with perfect hindsight what *could* have happened. The advantage of this approach, however, is to eliminate early on as many as possible of the problems identifiable from the viewpoint of each functional area element. Even if only half of the problems are identified in this manner, that is another set of problems that will not need to be addressed during the exercise.

EXERCISE DESIGN MUST BALANCE TWO KEY TRADEOFFS

Once the training objectives have been clearly spelled out and decisions made about the exercise mode, the exercise designers must strike a balance among several competing considerations. Two are key: freeplay versus scripting and the balance between the training audience and the training support personnel.

Freeplay Versus Scripting

The first tradeoff involves the balance between freeplay and scripted exercises. After describing the features associated with freeplay and scripted techniques, we compare the advantages and disadvantages of each.

In the broadest sense, scripted exercise events involve setting up a particular situation so that the training audience will have the opportunity to accomplish the desired training objectives. Scripted exercises take many forms, from rigid to dynamic. A rigid script presents a predefined set of situations to the training audience, regardless of the decisions leading up to those events. For example, if a script defines that on day three a breakthrough will occur in the southern sector, then regardless of the ability of the training audience to identify and counteract such an event, the event will still occur.

At the other end of the script spectrum stand dynamically scripted exercises. A good example of dynamic scripting is given by the German army's officer and staff training program. A corps commander, for example, wishes to train his subordinate division commanders in particular tasks. These tasks may be to identify the main enemy thrust on day one, execute a counterattack on day two, reconstitute a reserve on day three, respond to a flank threat on day four, and so on. He enlists the aid of his subordinate's brigade commanders to present certain types of situations to the division commanders. As the division commanders react to these situations, the corps commander decides to modify the situation further so that the division commanders' decisions have made a difference. At the same time, the corps commander, with the help of the brigade commanders, can guide the script dynamically so that the training objectives are still met.

Freeplay exercises, on the other hand, proceed without controller intervention from start to finish, and the sequence and type of situations presented to the training audience depend almost solely on the decisions of each side's commanders and the assessment of outcomes. The outcomes may be assessed by human umpires or by computers. Freeplay exercises may be guided somewhat by training objectives, but there is no way to guarantee that all, or even any, of the desired training objectives will be met during a pure freeplay exercise. However, a freeplay exercise may provide the most unexpected events, a situation that in its own way contributes to stressful training and provides a good diagnostic test of whether or not the force works well together.

Rigidly scripted exercises have the advantage that one can guarantee achievement of the desired training objectives or at least guarantee that the training audience will have the opportunity to perform the desired tasks. A well-designed scripted exercise can insert a rapid number of critical events or even events unexpected by the training audience and thereby provide stress in the exercise. Rigidly scripted exercises can contribute significantly to the process of teaching new tasks. However, such exercises have two drawbacks when applied to measure force proficiency.

The first is that it is terribly frustrating to the training audience that, no matter what actions they take, they cannot influence the course of events. Initiative receives no reward in a rigidly scripted exercise. The second problem is that the scripts themselves are rarely kept secret from the training audience, and therefore the exercises can be boring because nothing unexpected happens. The complexity of large-scale, and especially multinational, rigid exercise scripts means they are rarely created from scratch. The usual procedure is to pull last year's script out of the files and play it again, with only slight modifications. NATO had found, prior to the sudden political changes of 1989, that the scripted exercises looked the same year after year.

The German army has successfully used dynamic scripting. It provides a stressful environment in which the decisions of the training audience matter, and it ensures that the desired training objectives are met. The main reason that the dynamically scripted exercise works for the German army is that the senior commander is *the* trainer of his direct subordinates. The German commander does not tend to delegate the responsibility for training his subordinates to a staff officer. The senior commander defines and guides both the presentation of the threat and the overall assessment of results. Because the senior commander has already served in the positions of his subordinates, he has both the expertise to define credible results and the rank to enforce his assessments.

Dynamic scripting did not work so well in the past for the U.S. Army. The main reason appears to be that although the senior commander is doctrinally responsible for training his subordinates, the day-to-day job of training the subordinates was usually delegated to a staff officer responsible for training. Since the staff officer had neither the rank nor the experience to carry the weight of argument against the senior commander's direct subordinate, the staff officer was frequently overridden in training situations.

As a result, U.S. training branched in two directions: one direction was toward rigidly scripted exercises in which senior members of the training audience agreed ahead of time on the outcomes. (This branch became popular in NATO, as mentioned above.) The other direction was toward freeplay exercises in which the assessment of outcomes moved away from the staff officers. These assessments were provided either by umpires in the field for CFX and FTX, by manual look-up tables in CPX, or by a computer during CAX. Since freeplay exercises were more fun, more challenging, and more appealing to the drive to demonstrate initiative within the U.S. training audiences, freeplay exercises soon became the training mode of choice for the U.S. Army.

The advantage of dynamically scripted exercises is that they are likely to achieve the desired training objectives, provide sufficient leeway to allow the decisions of the training audience to matter, and can virtually guarantee the creation of a stressful environment for the training audience. The disadvantages of dynamically scripted exercises are that they: (a) require sufficient personnel with the proper expertise to run and modify the script dynamically and credibly; (b) are less likely to encourage initiative in subordinates, since results will probably be assessed to occur only within certain bounds so that the training objectives can be met; and (c) may not present the wide variety of extreme situations that could occur in wartime.

Freeplay exercises do encourage initiative, present unexpected events, may present a wider variety of the extreme situations that could occur in war, and can create very stressful environments. On the disadvantage side, they cannot guarantee any of the desired training objectives will be met; the initiative they encourage may lead to unsound practices; and their overall environment might even be not as stressful as dynamically scripted exercises, depending upon the course of events as they occur.

Of course, no freeplay exercise is completely free of scripted events, nor does a dynamically scripted exercise rule out elements of freeplay. The main question to be addressed is the desired balance between freeplay and scripted elements in a given exercise design. The degree of balance should be based, once again, on the training objectives.

If one is teaching new tasks, then scripting is probably the better type of exercise. Depending upon the tasks, conditions, and standards, one may choose a balance somewhere between the poles of rigid and dynamic scripting. The more repetition required to learn the new tasks, the more rigid the script should be. The less repetition required, the more leeway could be provided for the decisions and actions of the training audience to affect the outcomes and possibly subsequent events as well.

If one is evaluating previously learned tasks, then more flexibility is desired. For example, some events must be scripted to ensure that each training audience is given the opportunity to achieve the training objectives. Therefore, a significant amount of controller intervention and cooperation with the threat cells will be necessary to ensure that the training audience faces all of the necessary situations. At the same time, some freeplay must be allowed in this exercise so that the decisions and actions of the training audience are shown to matter. Usually, the scenario is designed so that it is unlikely that the training audience will completely wrest the initiative from the threat forces, thereby allowing the exercise control staff to maintain overall control of the flow of events through the cooperation of the threat cell.

Large-scale multiechelon training events, such as REFORGER, have favored freeplay over scripting. The training objectives have been broad to allow freeplay to flow in any direction, but this freedom has led to problems of the training objectives being neither specific nor achieved. At the same time, however, selected scripted events have occurred during large-scale freeplay exercises.

For example, the insertion of threat air assault forces into the rear area of the training audience occurred as a scripted event during CG 89. This insertion threatened the rear area, thereby forcing portions of the training audience to react. Due to the artificialities of scripting the insertion of these forces, many elements of the training audience were disturbed by the lack of realism. The intelligence staffs were given no warning, the air defense units did not have the opportunity to engage the forces on ingress or egress, and the RAOC was given no forewarning of the insertion. For these reasons, we recommended that this insertion be played as part of the freeplay exercise during CS 90, rather than scripted.

During CS 90, however, both freeplay air assault missions were so attrited that only a third of the penetrating force landed.[4] Freeplay was achieved, but at the cost of failing to meet one of the main training objectives.

There is no set rule as to how much of an exercise should be freeplay and how much should be scripted. However, the following three rules of thumb should help.

- Events necessary to achieve specific training objectives should be scripted in an exercise to ensure that the training objectives are met.
- Before the exercise, the training objectives should be prioritized by functional area and by echelon so that competing objectives and constraints may be overcome.
- The training objectives should be pretested so that competing objectives and constraints may be identified before the exercise.

In the first case, the training objective drives the need to script selected events. When inserting these scripted events into what is predominantly a freeplay exercise, care should be taken so that the scripted event does not detract from other training objectives (including exercise realism). For example, if a scripted insertion is supposed to take place, the intelligence cells should be given adequate chance to detect the event before it occurs, rather than having the insertion force appear magically behind friendly lines.

In the second case, it may be decided before the exercise that it is more important to train the RAOC to respond to a rear area threat than to train the intelligence staff to detect possible air assault assembly areas. In such a case, the exercise director should tell intelligence staff that this scripted event is about to occur and request its cooperation in training the RAOC.

The third case, pretesting the exercise design, has been discussed earlier.

---

[4]Part of the high attrition came about because for safety reasons the live helicopters in the exercise could not fly as low as they would in wartime.

Balancing Training Audience and Training Support Personnel

Selecting the balance between training audience and training support personnel depends upon many factors, including the chosen training modes. However, the selection of the training mode itself should be a function primarily of the training objectives and the needs of the training audience. The selection of the training mode and the required number of training support personnel should be secondary considerations when compared to the training objectives.

Exercises involve five groups of personnel: the training audience and four categories of training support personnel.

- The training audience, those personnel expected to achieve the training objectives.
- The exercise directing staff (di-staff) or exercise controllers.
- The evaluators of the training audience.
- The evaluators of the conduct of the exercise.
- The training mode support personnel, or the personnel required to make the selected training mode function properly (including the threat cell, if any).

The four categories of training support personnel tend to be double- or triple-hatted in small-scale exercises. During large-scale exercises, however, each person tends to focus on only a single training support category.

Training Audience. An example of the training audience in a small-scale exercise could be the maintenance battalion commander and staff. During a BCTP warfighter exercise, the training audience will probably be the commander and selected staff elements, with emphasis being placed on the G3 and G2 staff. For a large-scale exercise like REFORGER, the training audience could be the commander and his staff (G1 through G4) for all four echelons from battalion through corps.

Exercise Directing Staff. The directing staff (di-staff) for the exercise may consist of from one to thirty people or more, depending upon the size of the training audience and the training mode selected. The job of the di-staff is to ensure that sufficient opportunities are presented to the training audience so that its training objectives can be met. To accomplish this mission, the directing staff may need to intervene in an exercise to prevent certain situations from occurring and present different situations more in line with the training objectives. For example, it is unlikely that the training objectives will be accomplished if the training audience is either too successful or too unsuccessful against its opponents. A more balanced course of events tends to present better training opportunities and a more exciting and stressful training environment.

By its nature, a freeplay exercise tends to require more di-staff intervention because it is less likely that events will automatically proceed to satisfy the training objectives. The more scripted the exercise, the easier it is to control the sequence of events, and the less the directing staff must intervene in the sequence of events.

The threat cell should cooperate directly with the exercise directing staff to ensure that the training objectives are met. If the threat cell begins to play to win in a freeplay exercise, rather than focus primarily on providing training support, then many of the training objectives will not be met. Although some freeplay should be allowed for the threat cell, its major decisions should be cleared through the directing staff to ensure that they support the achievement of the training objective.

Evaluators of the Training Audience. During a small-scale exercise, the few personnel of the directing staff will also tend to take on the role of evaluators of the training audience and to lead the after-action reports (AARs). For a BCTP warfighter, in addition to the central exercise control staff, the observer/controllers tend to be primarily evaluators of the training audience but with some authority to influence local exercise control so that the training objectives of the staff element being evaluated may be met. For a large-scale multi-echelon exercise such as REFORGER, evaluation of the training audience may be performed by separate individuals without any exercise control authority. In large-scale exercises, however, the number of evaluators available tends to be insufficient to adequately measure each training element's achievement of its training objectives.

Evaluators of the Exercise. The evaluators of the conduct of the exercise are not interested so much in how well the training audience performed, but in how well the exercise itself provided the opportunities for the training audience to accomplish their training objectives. At all levels, the evaluators of the exercise tend to be personnel who are neither part of the training audience, di-staff, evaluators of the training audience, nor training support personnel required to make the training mode work properly. For a small-scale exercise, there may be no one actually evaluating the exercise per se, or that role may be taken by the next-higher commander or his representative. For large-scale exercises, such as a BCTP warfighter or a RE-

FORGER, the evaluators of the exercise may be from organizations independent of the trainers and the trainees, such as the RAND team asked to evaluate CG 89 and CS 90.

Training Mode Support Personnel. The last category of training support personnel is called training mode support personnel—those people required to make the selected training mode work properly. We use this term in order to distinguish them from the other categories of training support personnel. The roles that can be played by training mode support personnel include: the enemy threat; friendly adjacent, higher, or lower units; assessment personnel, such as umpires, referees, fire markers, or engineer obstacle controllers; or computer simulation operators.[5]

In a small-scale exercise, the role of the training mode support personnel may be taken by the di-staff. However, most training modes employed in large-scale exercises require additional personnel who are not part of the directing staff.

The number of training support personnel required depends upon the training mode selected. In FTX and CFX training modes, a large number of umpires is required to assess combat outcomes. In CPX training modes, training support personnel are required to control events and assess outcomes. In CAX training modes, personnel are required to cause the simulated units to act according to Red or Blue doctrine, depending upon the side being played; run the simulation and communications hardware and software; and assist the exercise controllers in the presentation of situations to the training audience.

One of the big drawbacks of CAX compared to other CPX training modes has been the large number of training support personnel it needs to move and fight simulated units. Almost every simulation designed for training purposes requires a person to move and fight every unit in the simulation.[6] The more units represented in the simulation, the more people are required. During CS 90, over 200 training support personnel were required to run the WPC simulations alone. The Army needs to pursue the development of automated or

---

[5]The function of the threat cell was described in the section on exercise directing staff.

[6]A notable exception is the Belgian IALTA model that defaults to automated "players" in the absence of human players. This is similar to the design of the RAND Strategy Assessment System (RSAS), a simulation designed primarily for analytic purposes using automated players, but allowing for human players to replace selected automated players. The Jet Propulsion Laboratory is currently examining ways to automate many of the decision processes in the CBS model.

semiautomated command and control for its training models to reduce the requirements for training mode support personnel.

Confusing Support Personnel with the Training Audience. During large-scale exercises, the distinction between training mode support personnel and the training audience frequently blurs. This tends to occur most often when there is a high ratio of training support personnel to training audience. When the ratio is high, the tendency is to redefine the training support personnel as part of the training audience. This creates three effects, only the first of which is desirable.

First, it does reduce the ratio of training support to training audience, at least in name.

Second, it raises the expectations of training mode support personnel that they will be trained as part of the training audience. Since no training objectives have been provided for them, no standards of evaluation are defined, and no personnel are assigned to evaluate or assess their training, these personnel are not usually trained in their doctrine. They may learn how to "game the game," but they tend to leave the exercise feeling frustrated that they were not trained. Their expectations were raised with no hope of being fulfilled.

Others, however, have considered their participation in such an exercise worthwhile. For many of the augmentees providing training support, it is the first time a junior officer or senior NCO has had to "put it all together" in the course of a complex operation. Factors such as air support, artillery tasking, task force creation, logistics considerations, and other factors all come together for the first time. Since this appears to be a frequent reaction, it may present an area where even a relatively simple computer simulation on a PC could be used to provide this type of synthesizing experience as a part of individual professional development.

Third, since the training mode support personnel are expecting to be part of the training audience, they are *not* expecting to be required to fulfill the duties of the training support personnel. The real purpose of the training mode support personnel is to act as a buffer between the training audience and the artificialities of the training mode.

In a CAX, for example, the computer displays are rarely in the form that would be available to the training audience in the field. Following the principle of "train as you fight," these displays should not be available to the training audience per se. The training audience should receive its information about the battlefield on the same type of equipment it would be using in the real world. The buffer between the training audience and the computer screen is the training mode support personnel. In addition to operating the computer, they present information to the training audience in a real-world form and by a real-world medium. If the coordinates are in hexes in the computer, they should be in universal transverse mercator (UTM) coordinates when presented to the training audience. If the map of the battlefield is in hexes in the computer, military overlays should be prepared for standard military maps.

Therefore, the role of the training mode support personnel is to reduce the effects of the artificialities of the training mode on the training audience. When their expectations are that they are there to handle this task, they accomplish this task. When their expectations are that they are the training audience, then they react to the training mode artificialities as would the training audience—negatively. Such players tend to learn how to "game the game" and play to win, rather than working to reduce the effects of the artificialities and provide better training support.

For example, when it was found that a data error allowed Blue headquarters and supply units to destroy Red maneuver battalions in CS 90, these headquarters and support units were used to hold the line and successfully prevent a Red breakthrough. In reality, headquarters and support units should not be able to defeat combat maneuver battalions. But even knowing this, the simulation players considered themselves part of the training audience (as they were sometimes told) and were therefore playing to win. Those who were frustrated by the artificialities of the simulation chose to take advantage of them rather than play in a more realistic manner and thereby provide better training support to the primary training audience.

The solution to this problem is to clearly distinguish between the training audience and the training mode support personnel. When this distinction is clear, then the responsibilities and goals of all personnel, training audience and training support personnel alike, can be carried out. Ambiguity in this distinction leads to inefficiency at best and failure to accomplish the training mission at worst.

One variation on this theme may be desirable due to the prioritization of the training objectives by functional area and echelon. Since it is unlikely that every training objective will be met in an exercise, ranking them will help determine which training objective has precedence in the case of a conflict. When a conflict arises, the training audience element with the lower-priority training objective temporarily becomes "training support" personnel for the element with the higher-priority objective.

For example, if the di-staff declares that a certain unit should have been moved south rather than north two hours ago to meet a corps G3 (operations) high-priority training objective, then the corps G2 (intelligence) cell will be responsible for buffering this artificiality for the G3 staff. The di-staff informs the assessment personnel that the unit will be "magically" moved to a southern location, at the same time informing the G2 staff of the decision and requesting their cooperation. The G2 then comes up with an explanation to present a realistic picture to the G3 staff. For example, the explanation could be that the enemy unit in question was undertaking a deception operation to appear to be moving north with part of its force, but actually the main force was moving south.

Temporary cooperation between the di-staff and selected elements of the training audience could make the achievement of the training objectives more feasible and the evaluation of the training audience more fair. For example, during both CG 89 and CS 90, magic moves done to satisfy the needs of the G3 staff and units in the field made it difficult for the intelligence staffs to accomplish their training objectives, frustrated as they were by the artificiality of the moves. If, instead, they had been informed of such events and asked for cooperation, their frustration would have been reduced and they could have been evaluated more fairly.

Overall, the ratio of training support personnel to training audience is a key tradeoff in exercise design. That ratio will often be close to one-to-one for large-scale multiechelon training events. This should not be a shock to anyone who has been involved in planning such an exercise. For every element in the training audience (e.g., G1 to G4, deep battle cell), there should be an evaluator of the training audience, and one or more training mode support personnel to ensure that the training audience is buffered from the artificialities of the training mode. In addition, there is also an exercise overhead of the directing staff, other training mode support personnel, and exercise evaluators. If there are shortages of qualified training support personnel, then the priorities defined for the training objectives by functional area and echelon can be applied to make best use of the personnel available. This prioritization applies to the employment of training audience evaluators and exercise evaluators, as well as to the training audience. When conflicts arise, selected elements of the training audience may be asked to cooperate with the directing staff for a short time to achieve higher-priority training objectives.

4. SUGGESTIONS FOR FUTURE EXERCISE DESIGNS

Our analysis of CG 89 and CS 90 leads to three major recommendations with respect to future large-scale, multiechelon exercises. First, given all the constraints faced and the increasing costs, most future large-scale exercises should consist of a single training mode, and the preferred training mode should be full simulation, as opposed to a combination of simulation and field exercises. To clarify this point, we recommend that the training audience operate from headquarters in the field through distributed wargaming, but no maneuver elements should be in the field when a combat simulation is the selected training mode.

Second, if combat simulations do become the primary training mode of future exercises, the simulations require substantial improvement before they are adequate to meet many of the training objectives.

Finally, due to the recent and dramatic world changes, future exercise designs should include, whenever possible, both active and reserve components, joint service elements, and forces from other nations.

USE FULL SIMULATIONS FOR FUTURE LARGE-SCALE, MULTIECHELON EXERCISES

Our analysis of the two test exercises clearly suggests that future exercises should not attempt to mix modes but rather should depend primarily on simulations. Using multiple training modes creates artificial seams that preclude the natural interactions between friendly and enemy forces. The numerous problems that arise when attempting to make the seams between training modes invisible detract from the achievement of the training objectives.

If someday there is a significant technological advance that actually allows a large number of simulated and live forces to interact in a truly seamless manner, then this recommendation will need to be reconsidered. In the meantime, because of the increasing constraints on maneuvering forces in field, the operational costs of using live heavy forces, and the recent advances in combat simulation technology, we recommend that simulations be used as the primary training mode for large-scale multiechelon exercises.

Centurion Shield 92 plans to employ simulation as the primary training mode, thereby significantly reducing both the cost and the complexity of the exercise. Costs will decrease for operational tempo and maneuver damage. The complexity of the exercise will be reduced by not including FTX, CFX, and CPX techniques in the same exercise. This is especially beneficial in that it eliminates the many interfaces that would be needed between each training mode.

Using simulation as the primary training mode also allows the threat to be represented by a group trained in threat doctrine. (For REFORGER 92, the representation will still be very "Red" in its representation.) This feature precludes the severe problems of Blue opponents facing each other on one part of the battlefield and facing Red opponents on other parts. It also precludes the confusing problem of having two overlapping Blue-on-Red and Red-on-Blue games over the same battlefield.

The recommendation to use simulations as the primary training mode does not mean that every type of training should use simulations rather than forces in the field. Some types of exercises are not adequately addressed by simulations, such as the deployment portion of the REFORGER exercise. Had it not been for the annual rehearsal of packing and moving large quantities of heavy forces long distance, it is unlikely that the U.S. Army would have been able to accomplish the deployment to Operation Desert Shield as well as it did. The "little things that go wrong" and slow the whole deployment can adequately be discovered only in a live deployment exercise. That is one reason why REFORGER 92 plans to include a draw of POMCUS assets as part of the deployment exercise to ensure that all of the little things work correctly. However, simulations could be used to analyze and train in the planning of such deployments and in training the higher staff elements in executing them.

Another area in which field exercises may be preferred over simulations is in logistics support for light infantry units. In the simulations, if the supplies are available, the receiving unit is assumed to have sufficient equipment to handle them. In reality, a light infantry unit has no way to handle pallet-sized loads of supplies. All supplies for light infantry units must be prepackaged into platoon and company packages. Otherwise, the unit cannot adequately distribute the supplies that may be on hand but not readily accessible. Issues of this nature will not tend to arise from simulations, but rather from field exercises.

From our observations we also conclude that to reduce the number of interface problems, the number of simulations of the same functional area should be kept to a minimum. For example, the interface between the two ground combat simulations (GRWSIM and CBS) is very complicated and may not be worth the effort. Rather than putting continued effort into building an interface between these two simulations, we recommend that only one simulation be employed in future REFORGERs. As part of the "Models Investigation Project" we performed for the Warrior Preparation Center, we recommended that the WPC substitute the CBS model for the GRWSIM model. The prerequisites for this transition included tying the WPC's air combat model (AWSIM) to CBS, expanding the CBS playbox to ensure that it is sufficiently large to handle the rear and deep operations, and ensuring that the CBS model is able to use the WPC's existing player interface. All of these prerequisites for CBS are currently funded and under way, and they may be completed by REFORGER 92. This transition will also provide the WPC with the corps-level model currently used by all U.S. corps, the United Kingdom forces, and possibly additional NATO allies. (See Allen, forthcoming.)

If this recommendation is accepted, we see the evolution of the experimental exercises as shown in Table 3. CG 89 had four training modes: FTX, CFX, and two simulations. CS 90 had primarily three training modes (except for light infantry): CFX and two simulations. REFORGER 92 is currently planned to include two simulations. We predict that future REFORGERs will include only one ground combat simulation.

Table 3

Training Modes in the Experimental Exercises

	CG 89	CS 90	REFORGER 92	REFORGER XX
FTX	X			
CFX	X	X		
Simulation 2	X	X	X	
Simulation 1	X	X	X	X

COMBAT SIMULATIONS REQUIRE IMPROVEMENT

Simulations do show promise in their suitability to train all higher-echelon (above brigade) functional areas. However, current simulations have many severe shortcomings in the way they represent different functional areas.[1] Broad areas that need improvement include the representation of combined arms effects, many different types of battles, key aspects of operational art, friction and fratricide, intelligence functions and products, electronic warfare effects, and the interactions between counterbattery sensors and shooters on each side.

Simulations Tend To Favor Armored Forces and Ignore Many Combined Arms Effects

Almost all simulations favor heavy armored formations over infantry, regardless of the situation, by ignoring many basic "combined arms" effects.[2] Combined arms have two types of effects: different branches of the three primary combat arms (armor, infantry, and artillery) are better suited to certain combat situations than others; and a mix of combined arms in each of these categories is essential in most frequently encountered combat situations. The following examples show that neither aspect of combined arms effects is well represented in most combat simulations, even though these are first-order or primary effects with respect to operational and tactical planning and assessment.

Most models apply a terrain multiplier that enhances either the lethality or the survivability (or both) of the defending unit if it is in rough, mountainous, or urban terrain. This multiplier tends to be independent of the type of unit, whether it is armor or infantry, which contradicts basic combat doctrine. In reality, infantry is more effective than armor at defending urban or mountainous terrain, but almost no current combat situations account for this first-order effect. For example, German II Corps participated in a simulation exercise in which its infantry, armed with sufficient antitank weapons and defending in prepared positions with minefields in mountainous terrain, was quickly overrun by threat tank battalions. This was blatantly unrealistic, yet many model proponents did not consider this lack of adequate representation a shortcoming that needed any modification. The CG 89 and CS 90 exercises were performed with these same basic limitations.

---

[1]This section will focus on simulations of mid- to high-intensity conflict. There is a related but different set of issues required for simulations of low-intensity conflict and operations short of war, but that set is beyond the scope of this report.

[2]CBS with the COBRA addendum does account for combined arms effects, as described below.

One could include a different combat multiplier for each type of unit for each type of terrain. This is a relatively easy and straightforward solution that could significantly improve existing combat models. However, such multipliers would be only the first step in representing combined arms effects. When units suffer attrition, they do not tend to lose combat assets proportional to the fraction of the force or strength represented by that asset. Historically, armor loss rates tend to be much higher than infantry loss rates, which in turn tend to be much higher than artillery loss rates. As a result, attrition can eliminate a unit's "combined arms" capability (in terms of containing an adequate number of assets from each branch) long before the unit has been destroyed. Therefore, a multiplier by unit type would be an inadequate representation of a unit's combined arms capabilities unless it was based upon actual asset holdings or at least the degree of attrition suffered by the unit.

At the moment, most models do not penalize a force for containing no infantry. As long as it has armor, it is assumed to be fully effective. Some models do not even let infantry assets contribute to the overall strength of the unit! This is in spite of the fact that infantry is the only asset on the battlefield that can take and hold terrain. Nor do simulations penalize a force with no artillery support, even if it is assaulting prepared defenses. If our combat simulations are to have any chance of providing realistic training support, they must correct some of these basic, fundamental errors.

In the case of the CBS model, the COBRA addendum has been created, which accounts for many of the basic combined arms effects described above. Fielded after the CS 90 exercise, COBRA is a rule-based submodel that runs on a computer separate from the CBS model. The COBRA addendum appears to account for most of the combined arms effects that need to be represented at that echelon. In addition, CBS has included an infantry infiltration methodology to better account for the unique abilities of light infantry. For more aggregate models, RAND has developed a methodology that accounts for combined arms effects in simulations that use aggregated combat assessment methods. This methodology is called "Situational Force Scoring," and it has been implemented in the RAND Strategy Assessment System (RSAS) (Allen, 1992). This is not the only possible methodology one could design to handle these first-order combined arms effects, but whatever methodologies are used, these shortcomings must be addressed.

Simulations Are Limited in the Types of Battles They Currently Represent

Many current models have difficulty representing certain types of battles, such as different phases of battle, river crossing operations, flank attacks, counterattacks, and passage of lines operations.

Phases of Battle. Many models represent only the assault phase but not breakthrough, exploitation, or pursuit. In the assault phase, a defender with a cohesive defense tends to have the advantage over the attacker. As a result, the attacker's loss rate tends to be higher than the defender's, thereby indicating an advantage to the defender. Historically, however, the defender has tended to lose more forces than the attacker. The reason is that if the attack was successful, the exchange rate shifted in favor of the attacker during the breakthrough, exploitation, and pursuit phases.

Unfortunately, many models represent only the assault phase, thereby always placing the attacker at a disadvantage. This has been compared to a continuous ambush in which the defender always kills more of the attacker. Furthermore, this representation actually models an infinitely elastic defender who is always cohesive and whose line never ruptures, regardless of how fast he is in retrograde. Unless our models consistently represent the different phases of battle, the models will be biased against the attacker and in favor of the defender.

River Crossing Operations. Some models have difficulty representing certain kinds of assault-type battles, such as river crossing operations. In most river crossing operations, for example, part of the force secures the opposite side of the river. This far-shore force is usually infantry-heavy, with limited antitank capability. This makes the force particularly vulnerable to artillery fire (before it has had time to dig in) and armored counterattack. Helicopters can help protect the far-shore force, but these assets tend to be susceptible to air defense fire. In addition, engineer bridging assets often are of no account in river crossing operations in some models. Few models account for all of these unique features of a river crossing operation.

Flank Attacks and Counterattacks. These two types of battles are discussed together since the problems associated with representing them in models are related. For these two types of battle, we will distinguish between the representations in aggregate low-resolution models and in detailed high-resolution models.

In aggregate models, a force that is performing a flank attack is given a force multiplier to reflect the benefit of that attack. Simple multipliers used for flank attacks tend to not exceed a factor of three, which thereby causes most counterattacks in simulations to fail. In reality, a flank counterattack by a division may be able to roll up the flanks of five divisions. This need not be represented by a very large force multiplier, but it could be represented by assessing a series of subbattles where each engagement is favorable to the flank counterattacking force. This approach better reflects the mechanism by which flank attacks and flank counterattacks succeed—by engaging pieces of the enemy under conditions very favorable to the counterattacking force as quickly as possible before the defender can prepare a cohesive defense. Unfortunately, many models allow the counterattacked force to shift quickly into at least a hasty defense, which tends to result in the counterattack being assessed as a failure.

In more detailed higher-resolution models, the problem of representing the benefits of flank attacks and counterattacks stems from too much information provided to each side and the ability to micromanage forces to achieve a cohesive defense quickly. Since the intelligence models tend to provide nearly ground truth, the players know almost immediately when the counterattack is occurring. Even when intelligence is limited, players are allowed to micromanage component units to achieve cohesive reaction quickly. As a result, a cohesive defense is prepared much too quickly in our simulations, when in fact the counterattacked force should have been virtually destroyed. It should not be surprising, with such shortcomings in combat simulations, that counterattacks against larger but flanked units do not appear to be worthwhile operations. (See the next subsection on operational art regarding the ability to micromanage forces without adequately representing the delays that would be involved.)

A recommended solution to the problem in higher-resolution models is to degrade the information available, delay the information over time, and delay the reactions of the threatened forces. All of these measures would contribute to a more realistic representation of these battles. These improvements can be implemented directly in the models, or one can try to ensure that the controllers enforce such delays on information and reaction. Due to the difficulties of ensuring these effects during an exercise, it is recommended that the models be changed so that the effects are implemented consistently for every exercise.

Passage of Lines Operations. Even though passage of lines is an operation rather than a type of battle, faulty representation of such operations occurs in models that assess no penalty for being engaged during the operation. During a passage of lines operation, whether advancing or withdrawing, both the passing and the standing force are more vulnerable to enemy action than they would be if only one of the forces were present. If combat were to occur during a passage of lines operation, both units would have degraded combat capability such that the two units in transition would not be as effective as either unit in place. However, combat simulations tend to represent units undergoing passage of lines operations as being stronger, simply because more force or combat assets are available in a given location. The fact that the two units are in an awkward posture for combat is not accounted for in these models. Therefore, the models present an unrealistic capability to swap fresh units on and off line under enemy pressure.

Overall Effect. If simulations do not represent the benefits of different types of battles, such as flank attacks, then no payoff accrues to employing good command and control. For example, if there is no significant benefit to a flank attack, then one is faced with only head-on assault options. As a result, many current combat models tend to ignore the benefits of maneuver and emphasize attrition effects. This situation also affects the representation of operational art, as described below.

Current Models Do Not Address Important Aspects of Operational Art

The most serious shortcomings of models are those that are the most difficult to quantify—the ones related to operational art. Basic model shortcomings related to operational art include the basic tradeoff between force, space, and time. For example, it takes longer for larger units to plan and execute cohesive maneuvers than smaller units, but most models do not account for this fact. If the model resolves units down to company size, then players are allowed to move each company in a division individually. As a result, divisional movement rates in these models tend to be approximately equal to the movement rates of companies, or about 20 km per hour sustained over a day. Actual divisional daily movement rates tend to average 10 km per hour or less. The model's overstatement has a significant effect on the representation of maneuver and subsequent combat intensity.[3]

[3]Hex-based models that represent only a single road from center of hex to center of hex encounter an additional problem in which movement rates are unrealistically slow. For example, a three-kilometer hex may represent an area that actually has two main

This ability to micromanage smaller units in a model to act more efficiently than the larger unit they comprise also affects combat assessment. During a flank counterattack, for example, the counter-attacked force can quickly posture its forces in a cohesive manner through micromanaging its smaller units, and therefore the counter-attacks tend to fail. Historically, a division-level surprise flank counterattack could roll up a force several times its size. The reason is that the counterattacking force is striking elements of the counter-attacked force sequentially, and it cannot react cohesively. Speed and shock are critical elements to the success of such a counterattack, but these are negated by allowing players to micromanage the force into a cohesive defense against a counterattack. The time delays associated with confusion over the actual situation and the ability to coordinate a cohesive defense are not represented in the models, thereby dooming most counterattacks to be assessed as failures.

The purpose of operational art is to balance forces, space, and time to achieve the objectives while fighting in as favorable a situation as possible. The failure to represent many basic aspects of operational art precludes our models from adequately training commanders and staff in the tradeoffs associated with operational art. The reason is that, even with an intelligence model, too much information about forces on both sides is available to the players, and the players coordinate the actions of many small forces without having to cope with the delays that would occur in planning and executing such a complex operation.

As an alternative to adding these necessary delay effects into the models, one may attempt to compensate for the model's artificialities by using training support personnel whose task is to delay information to the training audience and implement their orders in a more realistic time frame. This technique could go a long way to solving the problem as long as the training support personnel understood that they were not part of the training audience and comprehended realistic time frames for providing information and implementing orders. Once again, implementing these changes in the models will ensure the most consistent implementation of these corrections.

---

and four minor roads as having only a single road in the model, thereby creating much more congestion in the model than would be encountered in real life. One solution is to move away from hex-based models as the basis for representing road movement. Another solution is to represent multiple roads or increased road capacity across hexsides. In either case, the users need to compensate for the artificiality of a single road from center of hex to center of hex.

Current Models Tend To Ignore Friction and Fratricide

When a simulated unit is given a command, it follows those commands unless prevented from doing so by external factors. Such external factors may include enemy forces, enemy fires, friendly force density, or lack of supplies. However, many other internal factors not represented in our models affect a unit's ability to accomplish its mission.

For example, the unit or parts of it could get lost. The unit may not have sufficient time to plan a cohesive move, resulting in excessive congestion. The unit or its subordinates may find what appears to be an opportunity but turns out to be a delay. The right quantity but not the right mix of ammunition may be provided. All of these internal and many external factors can be grouped into what can be called "friction." To paraphrase Clausewitz, friction is the factor that makes even the simple things difficult.

Most of the preceding effects are not adequately represented in our current models. As a result, a commander playing in a simulation tends to plan that his subordinates will be able to perform at doctrinal norms, that movement rates will be near the average, that consumption will be near the planning factors, and so forth. If we continue to train our commanders and staffs in simulations without the penalties associated with not planning for the unexpected, we may be providing them with unrealistic training.

The advantage of field exercises is that many of these friction effects occur automatically, and thereby provide a good platform for realizing the benefits of planning for friction. If the Army is going to use simulations as the primary training mode, it is going to need to account better for friction so that proper planning procedures can be trained.

As suggested above, there may be ways to implement these corrections without changing the existing models. For example, inexperienced personnel operating the computer terminals tend to create their own form of friction through their lack of familiarity with the computer. However, personnel who are familiar with the simulations tend to cause forces to move in a "frictionless" environment. In addition, the Army's plan is to have the simulations operate on standard Army equipment, rather than on specialized terminals. Therefore, the friction caused by unfamiliarity with the computer terminals will be further reduced, and the forces will act again in a frictionless environment. Once again, the most consistent solution to the problem is to place the necessary improvements directly in the simulations.

Fratricide is a form of friction imposed by the friendly side. Attrition caused by friendly fire is both costly and demoralizing. Although it has occurred in every war, it has gained additional attention since Operation Desert Storm, where, due to an extraordinarily low overall casualty rate, friendly fire caused a high percentage of total U.S. casualties. As the concern for U.S. casualties continues to rise, the use of new technology and procedures to preclude fratricide will also tend to increase. As a result, our simulations will need to help personnel train with this technology and procedures, and penalize failure to follow the procedures with the risk of fratricide.

The key to modeling fratricide is to be aware that it is situation dependent. Losses caused by fratricide do not occur at a certain rate in a given type of battle but are a function of a number of qualitative factors, such as planning, training, and events preceding possible fratricide events. (For example, if the last six aircraft flying over bombed a unit, the unit is more likely to try to shoot down the seventh airplane, regardless of its identity.)

The simulated rate of fratricide should be a function of the degree of uncertainty presented to the shooters and not just a function of the type of unit, battle, or terrain. The higher the degree of uncertainty, the higher the likely fratricide rate.[4]

Current Models Do Not Represent the Intelligence Functions and Products Well

Many current models do not provide intelligence analysis training to the corps intelligence staffs. Most of the models provide too much information, thereby giving the players a picture that is nearly ground truth. Two primary factors contribute to this problem: units in our combat simulations tend to be represented as points rather than as areas, and the representation of uncertainty is limited in most models.

In the first case, many models represent units as single points without length or width. If a unit is detected, its location is known exactly to eight- or ten-digit coordinates. Even if the model represents the unit as a regular shape, such as a circle or a square of fixed size, it is relatively easy to determine exact unit location (such as the center of the circle or the corner of the box) for purposes of reporting or targeting in the model during an exercise.

---

[4]Contact the author for a briefing on suggested methods of modeling fratricide as a function of uncertainty.

In reality, units cover a wide area, and their locations for purposes of reporting are inexact at best. The location of even a battalion may cover many kilometers, depending upon posture. In addition, the shape of the unit will vary widely by situation, and attachments and detachments may occur. As a result, it is actually very difficult to report the location of real units, especially larger ones.

As long as our models represent units as points, lines, or areas of fixed shape and size, we will not be able to adequately represent the uncertainty involved in determining unit location. This is true for locating both enemy and friendly units, and it is an issue related to the fratricide discussion above. As a result, we recommend that the Army develop ground combat models for training purposes that represent units of varying shape and size that move on and interact with digitized terrain. This will go a long way in representing the uncertainty involved in determining friendly and enemy unit location.

In addition to improving unit representation in the models, the Army needs to improve the representation of intelligence processes and reports. One of the key model shortcomings is an inability to portray uncertainty. For example, in many combat training models, if a unit is detected, all the information about that unit is available to the opposing side. The opposing player may look up information on the unit including size, type, heading, speed, posture, identification, and even weapon holdings (depending upon the model). This detection process is binary—one either knows nothing about the unit or everything about it (even if that information is old).

Models need to represent the different degrees of knowledge about an enemy unit and how that knowledge evolves over time. In reality, one may know that something with tanks is located in this vicinity, but little else. One may later know the type and size of the unit, but not the direction of travel. One usually determines the identity of the unit last. Intelligence models supporting training exercises need to be able to present these levels of information, rather than the binary "all or nothing" information that many models offer. Unless this feature is accomplished, we will not be able to adequately represent the fog of war, nor will we be able to adequately represent deception operations.

For example, if a deception unit is attempting to pose as a brigade headquarters in place while the real headquarters is moving elsewhere, the intelligence model would have to represent the signatures of the deception unit as though it were a brigade headquarters. If only real units can generate enemy intelligence reports, then units or assets pretending to be something other than what they actually are cannot be represented in our models.

Current Models Do Not Address Electronic Warfare Effects

Most training simulations do not appear to represent electronic warfare (EW) effects except for air and air defense activities. Until the simulations significantly improve in this functional area, attempts to train through these modes should be kept to a minimum. This is a longstanding problem, due to the inherent complexity of electronic warfare. Both the WPC and CBS personnel have requested a definition from the ground EW and electronic combat organizations of what they expect to achieve, how, and under which conditions, so that the simulations can be improved in those directions. Both the long-term costs and benefits of this option are estimated to be high.

Current Models Do Not Train Counterbattery Fires Well

The representation of counterbattery (CB) fire is still very limited. In most models, the detection of artillery firing positions tends to be independent of any counterbattery detection systems or procedures. Also, any fires directed against an artillery unit detected as part of a CB program are also assessed against any other artillery unit that happens to be in the hex. Part of the difficulty in representing the employment of CB assets is that the interactions and procedures associated with these assets far exceed the simple presence of the assets themselves. In reality, a radar is fairly easily detected if it is radiating continuously. Therefore, the radar tends to transmit over short intervals. Similarly, tubes and launchers "shoot and scoot" to avoid receiving counterbattery fires. The representation of actual CB assets and procedures is much closer to a game of hide and seek between opposing assets. As a result, it is difficult to represent these interactions in simulations. Simulations need modification to allow explicit recognition of CB missions and a representation of CB interactions adequate to support training in this functional area.

Once again, the CBS model stands out as an exception to this generalization. Since the CG 89 and CS 90 exercises, the CBS model has obtained a reasonable and inexpensive (in terms of time and manpower required to operate) representation of counterbattery operations to support training.

Overall Observations About Model Limitations

The preceding limitations in current combat models need to be addressed to better support Army training requirements. Our analytical models share the same shortcomings. Unfortunately, although many of these problems persist across almost all simulations, many in the user community do not know that they exist. Proponents of a particular model tend to know its strengths but not its limitations. The complexity of detailed combat models disguises the fact that they do not accurately represent aggregate effects, such as division or corps movement and combat rates. Therefore, many model proponents assume that their models represent effects that in fact they do not.

Correcting all of the problems listed above poses a complex task. The National Simulation Center is the central clearinghouse of identified problems for most Army training models. Based on inputs from the user community, an Army panel or steering group meets and prepares a prioritized list of corrections. The priorities are a function of the training value expected from each correction and the time and money required to implement it. Unfortunately, the types of problems that *should* be considered first-order effects, such as adequate representations of combined arms effects, types of battles, friction, and passage of lines operations, are not usually raised to the top of the list by the user community.

Military officers understand how important these factors can be, but many do not realize that current simulations do not already account for these fundamental effects. When the representatives of the user community focus on tactical-level issues, the problems associated with operational-level issues are less likely to be identified. One goal of this report is to inform the action officers in the user community of the types of basic operational-level issues that simulations do not yet adequately address, in the hope that these first-order issues will be raised in priority for more rapid resolution.

Finally, when a model proponent declares that a given problem is "fixed," a knowledgeable representative from the user community that defined the problem needs to review the solution to ensure that the problem is indeed solved. This will help ensure that the problem was actually understood and subsequently solved.

EXERCISE PARTICIPATION SHOULD INCLUDE ACTIVE AND RESERVE COMPONENTS, JOINT ELEMENTS, AND OTHER NATIONS

Our final recommendations for future exercises are related to the recent fundamental changes in the international political and military context. The threat of a bipolar, high-intensity, conventional confrontation between the major powers has been significantly reduced. However, the threat of mid-intensity and low-intensity regional conflicts appears to have increased. Operation Desert Storm may well represent the types and combinations of friendly forces engaged in future conflicts.

For example, due to the drawdown of forces, it is likely that reserve and National Guard forces will be employed alongside active Army forces. Similarly, it is unlikely that the U.S. Army will be employed without close cooperation from the Air Force, Navy, and Marine forces. Finally, in light of the new political situation, it is less likely that U.S. forces will be employed unilaterally, but rather as part of a coalition of forces from allied nations.[5]

Each of the preceding combinations of forces presents unique training requirements and difficulties. This section will attempt to identify these requirements and problems and to suggest possible techniques for meeting our substantial and ongoing training requirements.

Active, Reserve, and National Guard Participation in Exercises

The Total Army program calls for reserve component units to serve alongside active units. A subset of the Total Army concept is the CAPSTONE program, which attempts to link reserve and National Guard units with specific active-duty units in case conflict does occur. This program encourages contact between the active and reserve/guard components in planning the training activities.

Because of the distances involved between these units when they are in garrison, it is difficult to coordinate joint training exercises between active and reserve/guard components. However, the distributed wargaming system (DWS) provides a new opportunity to simultaneously train CAPSTONE-linked units through simulation.

[5]As stated in FM 25-100, p. 1-1, "Moreover, for deterrence to be effective, potential enemies must perceive that the Army has the capability to mobilize, deploy, fight, and sustain combat operations in unified action with our sister services and allies."

During CS 90, a National Guard unit sent a 30-member staff element to participate in the exercise. Since this was a cohesive unit that had worked together before, it performed better than units manned by augmentees assembled solely for this exercise. In the case of CS 90, this element came over to Europe. However, using the DWS capability, the unit could have participated from its garrison, provided the communications link was established.

There is precedent for employing DWS to link training audiences on separate continents. During ACE 89, U.S. III Corps stationed in Texas participated in the NATO exercise via the satellite link of DWS. Although the III Corps personnel had to perform their duties during the night shift, the proof of concept was demonstrated.

As the active-duty component of the Army shrinks during the force drawdown, the role of reserve and National Guard units may continue to grow. The new technology of DWS will allow geographically dispersed units that may fight together to train together at comparatively little cost.

With regard to future REFORGER exercises, the size of the U.S. forces stationed in Europe will be reduced, but probably not to zero. The increased use of simulation as the primary training mode will create additional opportunities to have CONUS-based units participate in REFORGER exercises without the cost of overseas deployment.

The changes in Europe affect the training of not only U.S. forces, but also the forces of other nations. The last subsection discusses issues related to the participation of other nations in REFORGER and other exercises, the participation of multinational units in REFORGER exercises, and issues related to training multiple nationalities in a single exercise.

Joint Service Participation in Exercises

Although REFORGER has traditionally been an Army exercise, Air Force participation has been increasing. The results of Operation Desert Storm may further increase the cooperation between these services. The importance of joint service participation was observed primarily in air-to-ground coordination. Air Force liaison personnel assisted in ensuring that proper procedures were being followed in requesting air support. Air defense and air space management procedures were practiced. Forward Air Controllers practiced their procedures.

Without joint participation, the training audience may not be able to learn whether it performed certain tasks correctly. For example, during CG 89, a request came in to the Allied Tactical Operations Center (ATOC) for an air strike on specified coordinates. The ATOC personnel double-checked the coordinates, only to find that the strike was being called on a friendly city well within the rear area. It turned out that some digits had been transposed in the original message.

In another exercise, the Army and Air Force jointly focused on the creation of the BAI plan. Early in the exercise, the BAI plans were not well developed, but by the end of the exercise, they were prepared correctly. In another example, Forward Air Controllers were constrained by the arrival times of actual aircraft that had tightly limited times on station: if the target was not passed within that window, those sorties were lost. Even when one is using simulations rather than real aircraft, the presence of Air Force personnel knowledgeable on these real-world limitations helps add to the realism and training benefit.

Exercises should not be limited to just Army and Air Force participation. A change in focus from Central Europe to other parts of the world may make joint participation by airlift, sealift, and Navy and Marine personnel more important, even in primarily Army exercises. For example, both the Navy and Marines follow different procedures than the Army or Air Force and often use incompatible assets, such as communications equipment. Joint participation in exercises will help identify operational problems between the services before they emerge during combat operations.

Multinational Participation in Exercises

Although the REFORGER exercise is sponsored by the United States Army, Europe, the exercise has traditionally included units of other NATO nations. Since the Commander in Chief of USAREUR also serves as the commander of NATO's Central Army Group (CENTAG), most of the allied nations participating in REFORGER have tended to come from CENTAG. For example, the Canadian brigade and elements of German II and German III Corps have been frequent participants in REFORGER exercises. Participation of NATO allies in future REFORGER exercises is being planned as well.

One result of the reduction of U.S. forces in Europe is the need to include more units from other NATO nations. For example, the United States will not maintain two corps in Europe. It may be possible to play one U.S. corps staff from CONUS using the DWS capability described above, but the tendency will be to plan the basic exercise around forces already in Europe.

This likely future contributed to two fundamental features of the REFORGER 92 exercise. The first is that two NATO corps will not oppose each other in an exercise; rather, a simulated opposing force will be used against a NATO corps. (It is one thing to pit two U.S. corps against each other for purposes of an exercise, but to pit two corps from different nations against each other runs risks of offending political sensitivities.) The second is that the participation of an allied corps in the exercise is recommended for flank play. This feature will allow selected interactions with the Army Group headquarters to be practiced during this exercise, as well as selected functions associated with the coordination between adjacent corps. This exercise design has the added advantage of employing an actual corps headquarters as the flank corps, rather than a staff consisting of augmentees gathered only for the exercise.

In the event of a future armed conflict in Europe, it is unlikely that the United States will field as many "pure" U.S. forces as it did in the past. U.S. units will continue to fight together, but they may be attached to other NATO allied corps or have units attached to them from other NATO nations. It would be advisable, of course, to start training these forces and have them practice coordinating their activities now. For this reason, REFORGER 92 is designed to include U.S. units attached to the German corps, and German units attached to the U.S. corps. Other formations with mixed nationalities are also planned.

This mixing of nationalities in larger formations will help train coordination not only in the activities of the NATO G3 and G2 staffs, but also the G4 staffs. Since supplies are a national responsibility, procedures for the transportation of U.S. supplies in a German corps sector (or vice versa) must be practiced. These activities have occurred between selected NATO units in the past, but the upcoming exercises present the opportunity to standardize NATO-wide operating procedures.

One additional advantage of the shift to exercise multinational formations is that it helps NATO get away from the earlier "layer cake" force distribution along the border. The plan to deploy nine NATO corps adjacent to each other along the now-extinct inter-German border was politically attractive but posed a significant military disadvantage. This forward deployment of all corps made the forward defense of NATO particularly brittle and unlikely to withstand or recover from a major penetration. Furthermore, the distribution of force put the lightest NATO forces in the best tank country and the heaviest armored forces in the worst tank country. Finally, the need to mobilize and deploy NATO forces to their on-line corps sectors and replace the German border forces already in place contributed to a significant vulnerability. A "relief in place" operation is a difficult enough operation to perform, but doing so under the threat of an imminent and massive enemy assault increases the risk that the defense will crumble while still disorganized. Shifting the emphasis in future exercises to more multinational participation increases the flexibility of NATO forces to fight when and where needed under a wide variety of national mixes.

Multinational Unit Participation in Exercises. The issue of multinational exercises can be extended to multinational units. The ACE mobile force (AMF) has been the showpiece of NATO as an example of multinational cooperation. More recently, the Franco-German Brigade was created as an example of multinational cooperation beyond NATO. Within NATO itself, Northern Army Group undertook the creation of a multinational division to reduce the vulnerability resulting from the "layer cake" deployment scheme. Furthermore, the details for the creation of a truly multinational corps are in the making.

The multinational corps is designed to be a large-scale version of the multinational units described above and not just units of different nations attached to the corps of another nation. For example, rather than having a German division and a Canadian brigade attached to a U.S. corps, a multinational corps could have one division from one nation, one division from another nation, a separate brigade from a third, and a mixed multinational staff. A true multinational corps will be a significant step in the creation of a NATO-wide force that could function cohesively at any time and place. At the same time, it will be important for the United States to continue to participate in the creation and sustainment of such forces, since future conflicts may require our participation in these multinational formations.

In order to set a precedent and encourage the participation of multinational units in future exercises, we recommended the participation of the Franco-German Brigade in REFORGER 92.

Issues Involved in Training Multinational Forces Beyond NATO. The issue of multinational training goes well beyond the boundaries of NATO. Other U.S. forces forward deployed around the world have a similar need for multinational training. U.S. forces in Korea continue to have a need to train with Republic of Korea forces, and U.S. forces in South America have a need to train with the forces of the nations in that theater.

In addition to forward-deployed forces, U.S. contingency forces have a similar need for multinational training. The multinational forces arrayed against Iraq during Operation Desert Storm serve as an example of the degree of multinational training required. The United States was allied not only with traditional allied forces such as France and the United Kingdom, but also with Syrian and Egyptian forces that employed Soviet equipment and, to some degree, Soviet doctrine. To train together with such multinational forces requires a great deal of careful cooperation. In the absence of such cooperation, the likelihood of fratricide between allied nations increases substantially.

To help alleviate the problem of how to train multinational forces that include U.S. forward-deployed or contingency forces, *the United States needs to develop a multinational training doctrine.* This doctrine must be broad enough to handle the differences among a wide variety of functional areas but also specific enough to make the cooperation useful. For example, U.S. artillery doctrine differs significantly from German artillery doctrine. However, there are U.S. artillery missions that could be used to define which artillery missions German artillery will perform and which they will not. As a result, one could theoretically define a "superset" of activities by functional area in each echelon, of which each nation will be willing or able to undertake only a subset. The end result should be a standardized set of activities that each nation can agree to undertake in multinational operations; each can then train its forces to ensure that these activities work in a cohesive fashion. Future exercises should consider including such forces. Using simulations reduces both the cost and difficulty of doing so.

BIBLIOGRAPHY

Allen, Patrick D., *Situational Force Scoring: Accounting for Combined Arms Effects in Aggregate Ground Combat Models*, RAND, N-3423-NA, 1992.

Allen, Patrick D., *Evolution of Models at the Warrior Preparation Center*, RAND, R-4155-AF/A, forthcoming.

Allen, Patrick D., Jim Kahan, Robert Worley, Thomas Lippiatt, and Thomas Polsley, *Observations on the Centurion Shield 90 Exercise*, RAND, N-3152-A, 1992a.

Allen, Patrick D., Thomas Lippiatt, Leland Pleger, and Thomas Polsley, *Observations of the Caravan Guard 89 Exercise*, RAND, N-3151-A, 1992b.

Goldsmith, Martin, J. Hodges, and M. Burn, *Applying the National Training Center Experience: Artillery Targeting Accuracy,* RAND, N-2984-A, April 1990.

Headquarters, U.S. V Corps, *Caravan Guard 89 After Action Report (Draft)*, December 1989.

Headquarters, USAREUR and Seventh Army, *CINCUSAREUR REFORGER 90 After Action Report (4047x), Volume II, FTX Centurion Shield 90,* April 1990.

Kahan, James P., *Corps and Division Command Staff Turnover in the 1980s,* RAND, N-2994-A, October 1989.

Kahan, James P., D. Robert Worley, Suzanne M. Holroyd, Leland C. Pleger, and Cathleen Stasz, *Implementing the Battle Command Training Program,* RAND, R-3816-A, October 1989.