(12)

A RAND NOTE

Utilizing the Data from the Army's National Training Center: Analytical Plan

Robert A. Levine, James S. Hodges,
Martin Goldsmith

June 1986

RAND

The research described in this report was sponsored by the United States Army under Contract No. MDA903-86-C-0059.

Published by The RAND Corporation
1700 Main Street, P.O. Box 2138, Santa Monica, CA 90406-2138

A RAND NOTE

N-2461-A

Utilizing the Data from the Army's National Training Center: Analytical Plan

Robert A. Levine, James S. Hodges,
Martin Goldsmith

June 1986

Prepared for
The United States Army

RAND

PREFACE

The Arroyo Center project designed to assist the Army with "lessons learned" from its National Training Center (NTC) located at Ft. Irwin, California, is currently in its second phase.

The first phase, lasting from 1983 to 1985, had a dual purpose: to enable Arroyo Center analysts to familiarize themselves with the Army's *modus operandi* at NTC, and to analyze the ways in which the lessons being learned from NTC were spread throughout the Army, in order to improve their dissemination. This phase is reported on elsewhere.[1]

In the course of Phase I, it became apparent that, because the primary purpose of NTC's operations and data collection is the training of Army units rotated through the installation, many potential lessons from the unprecedented wealth of data being collected on the Center's realistic operations were being left unstated and unanalyzed. The data themselves were being underutilized and eventually lost. For this reason, in the spring of 1985 the focus was shifted to assisting the Army to derive lessons, using the vast quantity of data gathered by NTC's instrumentation and its skilled officer cadre.

Phase II also has two objectives:

- To *develop a methodology* with which the Army, with the assistance of the Arroyo Center, can use the materials from NTC to derive lessons on a continuing basis.
- To *derive exemplary lessons*, primarily but not exclusively in the area of doctrine, that will help the Army improve its combat capabilities.

This Note lays out the approach to the first objective, the development of a methodology. Other reports will discuss specific lessons (e.g., suggestions for lowering the rate of ground-to-ground

---

[1]Martin Goldsmith, *Capturing Lessons Learned: Applying the National Training Center Experience*, The Rand Corporation, N-2384-A (forthcoming).

fratricide) as they are drawn from NTC operations. It is intended to assist those planners, primarily within the Army, who want to know:

- Details of NTC operations.
- The kinds of data actually or potentially available there.
- What can be done with those data, through lessons learned, to help improve Army combat capabilities.

THE ARROYO CENTER

The Arroyo Center is the U.S. Army's Federally Funded Research and Development Center for studies and analysis operated by The RAND Corporation. The Arroyo Center provides the Army with objective, independent analytic research on major policy and management concerns, emphasizing mid- to long-term problems. Its research is carried out in five programs: Policy and Strategy Studies; Force Development and Employment; Army Readiness and Sustainability; Manpower, Personnel, and Performance; and Applied Technology.

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, 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 MDA903-86-C-0059.

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

Stephen M. Drezner is Vice President for the Army Research Division and Director of the Arroyo Center. Those interested in further information concerning the Arroyo Center should contact his office directly.

Stephen M. Drezner
The RAND Corporation
1700 Main Street
P. O. Box 2138
Santa Monica, California 90406-2138
Telephone: (213) 393-0411

SUMMARY

The Army's National Training Center (NTC) is a large (nearly 1000 square miles) maneuver range located at Ft. Irwin in the California Desert. It is unique in the quantity and variety of instrumentation that collects, records, and collates a wide range of data on realistic simulated force-on-force battles between U.S. Army units brought in for training, and a permanent "Opposing Force" made up of U.S. Army units simulating Soviet tactics and weapons.

The Arroyo Center NTC project is designed to create a methodology for the Army to derive "lessons learned" from the NTC, using data from the NTC's instrumentation and its officers, or gathered by Arroyo or other analysts. There are 17 types of data by one mode of classification. This methodology is being developed inductively and empirically, by collecting and using NTC data for policy analysis to help the Army solve specific problems bearing on combat capabilities.

The methodology has four steps:

- *Choice of issues.* Thus far, work has taken place on two issues: fratricide and utilization of artillery at brigade and battalion levels. Both are important to the Army and both can be illuminated by NTC data and analysis. Each, however, makes use of only a narrow portion of the range of available data, and the next problem will be chosen to assist in broader methodological development.
- *Hypothesis formulation.* This crucial analytical step draws on knowledgeable NTC and other personnel to:
    a. Describe the issue and situation operationally, in terms of concrete observable actors, relationships, and events.
    b. Divide the data--measures of the events just defined--among dependent variables, policy variables, and other independent variables.

c. Formulate tentative hypotheses from the situation description and the available data.

d. Recycle and refine the hypotheses through expert discussion.

As an example, these steps have been applied to the use of reconnaissance and scouting, and their importance to battle success at NTC.

- *Data selection and collection*. For purposes of analysis, NTC data can be classified in several ways according to their quantifiability, the degree to which they have been predigested and preanalyzed before becoming available for the particular problem at hand, and the amount of special planning needed to collect them. For any given issue, preparations must be made to retain or gather the relevant data.
- *Hypothesis "testing"*. "Testing" is in quotes because the real world never produces conditions or data sufficient to test hypotheses in any rigorous logical sense. Rather, what is done is to move again through the steps of the Hypothesis Formulation phase, again using Army expertise, to compare data and data analyses with the knowledge of experts. Where the analyses and the expertise agree, it may be possible to make strong doctrinal or other recommendations to the Army. Most officers believe, for example, that good reconnaissance has a very high correlation with battle mission success. Should the data analysis confirm this, recommendations may be forthcoming concerning the stress put on reconnaissance during training. Where analyses and expertise disagree, however, one need not be chosen over the other--certainly not analysis of imperfect data over trained judgment. Rather, further analysis and discussion should lead to understanding of the differences and to both improved analysis and perhaps revision of the expert judgment. It may be, for example, that reconnaissance is not equally crucial to *all* types of engagements, and discussion of whether and why this is the case can be more helpful to the Army than any cut and dried "answer".

CONTENTS

I. INTRODUCTION

The U.S. Army's National Training Center (NTC) is a very large (almost 1000 square miles) maneuver range at Fort Irwin in California's Mojave Desert. It is dedicated to the simulated-battle training of battalion-size units from the heavy divisions and independent brigades of the Forces Command (FORSCOM) stationed in the continental United States. The training program is the responsibility of the Training and Doctrine Command (TRADOC).

NTC is unique in two ways.

- It is highly instrumented, both with the Multiple Integrated Laser Engagement System (MILES) which allows realistic simulation of weapons engagement and with range instrumentation which follows the location of most vehicles and some dismounted infantry forces. This instrumentation makes possible the collation, real-time display, and recording of a wide variety of battle data.
- It has a standing Opposing Force (OPFOR) which uses Soviet tactics and simulated Soviet weapons in force-on-force engagements with the FORSCOM units being trained.

These two features make possible battle simulations far more realistic than any others carried out herefore. Further, an Operations Group of Observer-Controllers (OCs) responsible to TRADOC manipulates the engagements to provide maximum realism and training value, and the OCs gather their own data for post-battle analysis. The addition of these observations to the instrumentation data provides a rich base of qualitative as well as quantitative data for potential systematic analysis of these closest-to-the-real-thing exercises.

The development of a methodology for utilization of the data in this kind of analysis is the primary purpose of Rand's Arroyo Center NTC project. This Note sets forth the plan for the methodological development. After describing the operations of the NTC and the data

gathered, the Note describes the policy analysis and statistical methods to be employed, and then the ways in which the project intends to apply these methods.

II. NTC OPERATIONS AND DATA

NTC OPERATIONS

An NTC training rotation consists of the deployment to Fort Irwin for 14-18 days of force-on-force engagements (and separate live-fire exercises, which are not covered here) of two battalions, generally one mechanized infantry and one tank, from a FORSCOM heavy division or separate brigade. The division sends in addition a brigade slice of assets, including the brigade HQ, a forward support battalion, elements of the division artillery, an appropriate share of the division air assets, and combat engineering support. On occasion, a cavalry squadron will substitute for one battalion, or the heavy forces may be supplemented by a light infantry battalion. Close air support is provided by the U.S. Air Force during battle engagements.

The FORSCOM Blue Forces (BLUEFOR) are trained against a standing Opposing Force (OPFOR), which consists of two U.S. heavy battalions, one infantry and one tank. The infantry battalion lacks most of its dismounted element, which is often provided by other U.S. infantry units. The two battalions are configured for battle as a motorized rifle regiment of the "Krasnovian" forces, and replicate a Soviet-supplied Warsaw Pact unit. For the most part, their vehicles are U.S. equipment visually modified to resemble T-72 tanks, BMP and BRDM personnel carriers, etc. Their battle doctrine is that attributed to Pact forces.

The training is conducted by teams of observer/controllers (OCs), who are U.S. Army officers and NCOs on regular assignment to NTC. A team of about 30 or more is assigned to each training battalion, and accompanies the BLUFOR throughout their rotation. The team's function is to control the battle, assess results, and provide an After Action Review at the conclusion of each engagement.

The conduct of the training is assisted by the Core Instrumentation Subsystem (CIS), a notable feature of the NTC. The training area at Ft. Irwin, which consists of a large expanse of the California high desert (over 600,000 acres), incorporates a series of radio position/location (p/l) stations. These stations communicate with p/l units installed on

BLUFOR and OPFOR combat vehicles, and carried by some dismounted infantry units and some observer/controller vehicles. By triangulation, the position of each vehicle can be determined by the CIS, which provides data to a central Training Analysis and Feedback (TAF) facility near Fort Irwin post headquarters. This information is displayed at a TAF operations center on graphics terminals where the operational information is superimposed on map displays of the training area with a choice of scales and cartographic backgrounds. These displays can also show overlay graphics prepared for the battles as part of the planning and orders process. Thus an analyst stationed at one of the terminals may observe the position of the engaged units during the battle, and moreover can replay prior action at any time during or after the battle.

In the force-on-force exercises at the NTC, MILES is used to simulate weapon engagement. Each direct-fire weapon system (i.e., a system aimed along a line of sight to a target, as compared to artillery fired on a high trajectory) is equipped with an eye-safe laser boresighted to the weapon. When the weapon is fired (with blank ammunition or a simulator), a coded laser beam is emitted. Each individual player and each tactical vehicle is equipped with laser receivers which register hits by the laser designators. If a soldier is hit by an M16 rifle code, his MILES set will register the hit with a piercing audio tone, which will indicate to all that he is a casualty. On the other hand, if a tank registers a hit by an M16 code, nothing happens, because a rifle cannot kill a tank.

When a tank main gun fires, several things take place. The coded laser beam is directed at the target, a simulator charge is fired that yields a visible and audible signal, and a firing message is sent through the p/l unit to the CIS. Should the laser beam hit a target vehicle squarely (kill probabilities can be accounted for), the target's instruments will register the code of the weapon type, disable its firing mechanism if it is a tank, start an externally mounted strobe light, and send a kill signal to the CIS. The CIS, on receiving such a signal, will search for a firing message to match in character and time; when one is found, a pairing is made. (Frequently, however, a pairing cannot be made owing to signal masking or other instrumentation problems.) The graphic display will then show a firing vector between

the units, and a kill, if that is the result of the hit. The instrumentation system keeps a record of the near misses, hits, and kills, shows the locations of the firer and target, calculates the range, and keeps cumulative scores.

A deficiency of the system is that firing of infantry weapons, such as rifles, Vipers, or dismounted TOWs, is not recorded, and therefore paired kills from such sources cannot be seen; the kill will simply be recorded as of unknown origin (although the original field data stream contains the killing weapon type). Moreover, MILES cannot be used with indirect-fire systems, so the simulation of artillery and mortars still involves subjective assessments. At the present time, the helicopters are equipped with MILES, but are not linked into the CIS. Fixed-wing aircraft are not equipped with instrumentation, and close air support and air defense are played in a subjective fashion. Thus the CIS battle record is of great value, but is by no means complete.

When a unit arrives at the NTC, its command group is issued orders from a fictional division HQ which establishes the upcoming mission and lays out the situation in terms of neighboring friendly forces (notional) and the OPFOR. The scenario usually involves a Krasnovian invasion of a U.S. ally, Mojave. The task force command groups then begin their deployment to initial positions, and prepare operations orders. As the battles progress, new situational information is issued by the divisional HQ (actually a section of the NTC Operations Group). As each battle is terminated by the Ops Group, the task forces must undertake the problems of real-world repairs and resupply, plus such simulated efforts as evacuation, reconstitution, and ammunition resupply, all of which must be accomplished in real time by real assets.

Throughout their time at the NTC, the training units are closely observed by the controller teams. At the conclusion of each battle, the company and platoon level observers conduct After Action Reviews (AARs) in the field. The battalion level AAR occurs a few hours after the battle, and involves the battalion staff and commanders. This review is limited to two hours, and takes place in a mobile TV van located near the battle area. Equipment in the van is capable of displaying the CIS graphics, which help illustrate important points about the day's action. The interactive AAR process is videotaped for future use by the training

unit. This tape, plus written summaries of the notes made in the field by the observer/controllers, are part of a take-home package supplied to the unit at the end of its rotation.

NTC DATA

As noted above, NTC data may be gathered in two ways, by electronic or by human means. Much of the electronic data is transmitted electronically to the TAF facility; most but not all of it is processed and recorded, sometimes in a rather elaborate form. To create an initial checklist, it is convenient to use the electronic/human division to categorize the data types, although, as will be seen, the most useful data classification for purposes of analysis is somewhat different. It should be noted that, although all of the data types listed here are in fact gathered and/or recorded, not all are preserved for very long. The first three items listed--types of data which are gathered, transmitted and processed electronically--are preserved, and make possible the "replay" of an entire battle or selected portions on graphic and computer consoles, either on a real-time scale or in speeded-up mode. Supplementation of these data types with others not handled systematically, however, would greatly facilitate analysis, and the Arroyo Center analytical plan calls for much more regular gathering and preservation of these other types of data than is presently the case.

The types of data actually or potentially available from NTC are:

Data gathered, transmitted, and processed electronically

1. *Graphics*, which combine position-location data, firing data, kill data computed or entered manually to record kills assessed by OCs, terrain data, and planning and other data pre-entered to provide to the TAF operations centers dynamic graphic/cartographic pictures of the battles as they take place, with a complete record available for replay.
2. *Real-time digital data*, which provide, on screens next to the graphic displays, a digital record of fires and kills as they take place, also with a complete record available for replay.

3. *Computer summaries*, cumulative records available as the action takes place, and at the end of the action, for a large number of statistics, ranging from kills and fratricides to numbers of transmissions by time over various communications nets.

Data gathered and transmitted electronically, but not easily retrieved

4. *Unprocessed data*, gathered and transmitted to the TAF via the CIS, but not further used. An example here includes some kill codes, by which MILES records the type of weapon that killed a particular vehicle. The kill code is transmitted to the TAF, but unless the pairing system has computed the specific killer as well as the victim, it is not easily retrieved. It could, however, be built into the retrieval system. The Army is considering this and the features.

Data recorded electronically, but not transmitted to the computer system

5. *Communications tapes*--recordings of all audio transmissions on up to 40 communications channels. These are retained by the Army Research Institute, but no current system allows them to be synchronized automatically with the graphics to form a coordinated record of what is being said at the same time as action is taking place. (Such a synchronization system could be developed in the future for NTC data analysis.)
6. *Video tapes* of battle segments televised by cameras placed on various high points on the range. Much of the maneuver ground is covered by video, but those parts of the engagement that take place in the dark are not available. The videos are useful mainly to give an overall view of such conditions as dust and smoke.
7. *After-Action Review (AAR) videos* made of all task-force level AARs, which take place a few hours after each engagement, and of all summary AARs at the end of rotations. They are also available for some special (e.g., artillery, materiel) AARs, and seldom for company and lower level reviews.

Data gathered non-electronically, but relatively systematically

8. *Scenarios*, the Operations Group's plans laying out the evolution of the entire rotation--what kinds of battles will take place when and where. For example, that the day after the armored battalion attacks the OPFOR under specified circumstances, the mechanized battalion will defend a particular geographic segment against OPFOR counterattack.
9. *Operations orders*, both BLUEFOR and OPFOR, the specific, usually written orders given to subordinates before each engagement by the commanders of the BLUEFOR task forces and of the OPFOR. Frequently supplementing these pre-battle orders are fragmentary orders issued, usually orally, during the battle ("fragos"). These ordinarily exist only on the audio recordings of the communications nets (see #5 above).
10. *Take-home packages*, the final summary analyses of an entire rotation provided to the rotating unit. These are readily available but of limited utility for analytical purposes, since they summarize a wide range of relevant information in a very short compass.
11. *OC kill and related records* supplement the instrumented data on kills. The details of a kill are recorded in the CIS only when the computer program is able to pair a firing and a hit. In the field, however, MILES records a kill and the type of weapon doing the killing whenever the firing laser hits a target capable of being killed by the weapon fired; and the killed vehicle records the weapon type of the killer. OCs list all kills, paired and unpaired, to complete the killing record.
12. *Artillery logs* are the detailed records kept by the officers of the artillery TAF. Since indirect fire cannot be recorded by MILES, the firing of artillery and the area of impact are entered manually into the computer system by analysts in the TAF, forming part of the graphic record. The results of the impact, also assessed in the artillery TAF, are not entered into the computer. Rather, these and all summary data are kept in manually maintained logs in the TAF.

13. *Special-purpose interviews* are occasionally obtained, but this practice is limited by the work load of the Operations Group. One recent example of such interviews to gather OC and other observations comparing similar weapon systems. On a more continuing basis, it has been proposed to the Army Research Institute (ARI) that leadership lessons be distilled from a carefully designed set of interviews.[1] As will be noted below, this technique may be quite useful for the kinds of analysis proposed in this Note.
14. *Analyst observations* are observations planned to gather data relevant to the issue at hand, gathered by individuals assigned to the analytical function, including Arroyo Center analyst/observers as well as those officers assigned to the NTC by the Combined Arms Training Activity (CATA) of TRADOC. Unlike the observations of the officers of rotating units and NTC cadre, which are incidental to their primary duties (e.g., fighting battles for BLUEFOR and OPFOR officers, controlling and training for OCs), these observations are designed to provide specific information for the particular problem under analysis.
15. *Pre- and post-NTC data* from the rotating units. Examples here include materiel records: What is the state of their equipment and supply going into NTC (presumed to be as high as it can be made by the rotating units); what is their state leaving NTC and how long does it take for restoration to a new battle-ready state? CATA presently conducts post-rotation interviews on a regular basis.

Non-systematic data

The data listed here are the "softest" and the least quantitative, but as will be discussed below, are potentially the most valuable of all the data gathered at NTC.

---

[1]Donald R. Jones and Kerm Henrikson, *Development of a Leader Lessons Learned Methodology*, Allen Corporation of America, Report submitted to the U.S. Army Research Institute, October 1985.

16. *OC notes* are the observations collected by the observer/controllers for use in AARs and other summary battle analyses. They fill a major gap left by the more "regular" methods of data gathering, particularly the electronic ones; they apply expert human intelligence to the question of "What's important?"
17. *Trainer observations* are the more generalized observations of personnel in various field or TAF positions--OCs, BLUEFOR, OPFOR, analysts in the TAF. Most of the officers and NCOs involved in NTC activities are observant and analytically minded even if not trained in formal analysis, and remarkably articulate and voluble. This unparalleled data source is available, and when approached systematically and tactfully can provide a wealth of information and ideas not available in any other way. These observations are presently being utilized in lessons learned products.

ADVANTAGES AND DISADVANTAGES OF NTC

The wealth of data is one of the unique advantages NTC has as a source of doctrinal and other lessons for the Army, but it is not the only advantage. The entire list includes:

- The data.
- The fact that running similar battle scenarios on similar terrain over and over again allows a closer approach to replication than is possible in any other way. (El Alamein was unique. NTC in a sense runs several Alameins a year.)
- The graphics and related systems that make it possible to reexamine as frequently as necessary the events in any segment of any battle.
- Probably most important, the carrying out of realistic force-on-force exercises by U.S. Army battalions against other battalions simulating potential opposing force (OPFOR) weapons and tactics.

Against these advantages must be placed a number of disadvantages:

- Shortcomings in data and instrumentation.
- The fact that the pressures of battle cannot be fully reproduced in any training exercise.
- The related fact that safety constraints must be imposed.
- The terrain and other external conditions do not match many of the locales in which a future war may take place.

Many of these shortcomings may be minimized: by careful adjustment (e.g., by making sure, insofar as possible, that reported data present an unbiased sample of unreported data and by correcting for the remaining biases); by making specific allowances (for example, the existing practice of requiring units to exercise medical and replacement procedures as if casualties were real); and by focusing analysis on issues where simulation shortcomings are least crucial (e.g., command capabilities to readjust plans as the situation changes). Sometimes the shortcomings will predominate and preclude the use of NTC data and experience on certain issues (e.g., until better instrumentation is placed on fixed-wing aircraft, NTC data are close to useless for analysis of close air support). In any case, however, awareness of the shortcomings, compensating for them insofar as possible, and modesty about the meaning of "findings" can together preserve the unique advantages of NTC data and operations for the analytical derivation of doctrinal and related lessons for the Army.

III. USING NTC TO DERIVE "LESSONS LEARNED"

POLICY ANALYSIS AND STATISTICS

The frame of reference for development of lessons that can be applied by the Army and used by its soldiers is that of *policy analysis*. Policy analysis is directed at providing "How to" options to the Army, i.e., systematic alternatives for action. An example is, "Revise the methods of calling in artillery fire in the following specific ways." It contrasts with "purely descriptive" analysis, which might discover that artillery fire is ineffective compared to some standard in certain kinds of battles, but cannot suggest that such battles be avoided for that reason. "Descriptive" analysis can provide an important step toward "How to," but for the Army to benefit from NTC lessons, the firm objective must be the "How to" type of analysis.

Like any other analytical problem, the derivation of lessons from NTC can be divided into three phases: *description* of the overall situation (sometimes called "modeling"); *generation* of hypotheses about Army doctrine and related issues, in that situation; and "*testing*" of the hypotheses using NTC data.

- The necessary first step is to describe the problem and the context in which it arises. Such description includes (i) identification of the relevant factors (e.g., the forces and their subunits, visibility); (ii) evaluation of potential states of these factors (the OPFOR may be at full strength or at a specified lesser level for a particular engagement, visibility may be good or may be obscured by blowing dust); (iii) expected causal relationships (Soviet tactics call for specified changes in case of understrength defenses, substantial dust reduces effective visual ranges by x meters); and (iv) possible projections from the combination of i, ii, and iii (e.g., the dust made it difficult for BLUEFOR to know that OPFOR was defending understrength and the outcome depended on the skill of Blue scouts). The need for such description

may seem obvious, but in fact much data analysis suffers from its omission and becomes impossible to translate into policy analysis; in the absence of such explicit description, the data, hypotheses, and conclusions--bereft of a context--cannot be generalized to other situations. This is particularly important for NTC because its activities *are* simulated, and the extent to which deficiencies in simulation may introduce bias into results is always an issue in the interpretation of NTC data.

- A hypothesis is simply a hunch about the way things work--what might cause what; what might cure what; what might avoid what; what might improve what. Generation of hypotheses requires no analytical rigor. A hypothesis may come from anywhere--from Observer/Controllers (OCs) or other soldiers at NTC, from Army personnel elsewhere, from Arroyo analysts working with NTC and NTC data. Patterns in the data themselves may suggest hypotheses. The next subsection of this Note discusses ways in which the resources available at and to NTC may be organized to systematize hypothesis generation and bring the most promising hypotheses to the stage of "testing."
- "Testing" is put in quotes because neither at NTC nor almost anywhere else in the real world is formal "proof" of a hypothesis possible, not even in the statistician's sense of "failure to reject a null (i.e., converse) hypothesis." The data shortcomings and artificialities of NTC mean that the analytical product will be quite far from what passes as proof even in some of the social sciences. Proof of battle doctrine comes in battle only, and, while the near-real battles of NTC can provide near-proofs (with proper caveats), such near-proofs cannot automatically supersede informed judgments and intuitions of experienced officers. Indeed, in the testing process proposed here, some of the most valuable outcomes stem from the cases where the results of data-based tests *differ* from intuition; figuring out why they differ can help illuminate the real issues at stake.

Our proposed scheme for "testing" hypotheses has three parts, the second of which corresponds to the traditional notion of a formal hypothesis test. The first part consists of describing (i.e. reducing) the data in some manner appropriate to the hypotheses of interest; in the second, a determination is made about whether observed differences are "big" differences; the third involves the attribution of observed "big" differences to a cause or causes.

Data Description

"Data description" (or reduction) is a relatively new approach to data analysis developed, most notably, at Bell Laboratories.[1] The object is to find a cogent and succinct description of the data without invoking stochastic (i.e., formal statistical) models. In doing so, it is possible to use:

- Display techniques, to look (for example) for trends in battle data or conversely for unusual battles.
- Search techniques such as the Classification and Regression Tree (CART)[2] which searches in a group of variables for the ones that best differentiate between two classes of items, say, battles with some event of interest and those without.
- Various sorts of regression methods, where the object is to describe trends and relationships between variables without claiming any causal relationship or postulating any probabilistic model.

Of course, a hypothesis may suggest an obvious description and remove the need for such techniques.

---

[1]See Colin L. Mallows, "Data Description," in G.E.P. Box, T. Leonard, and C.-F. Wu (eds.), *Scientific Inference, Data Analysis, and Robustness*, Academic Press, New York, 1983.

[2]See Leo Breiman, Jerome H. Friedman, Richard A. Olshen, and Charles J. Stone, *Classification and Regression Trees*, Wadsworth, Belmont, California, 1984.

Determining Bigness

In statistics, the purpose of a hypothesis test is to decide whether some observed difference is "big"--i.e., big enough to appear unlikely unless the hypothesis being tested is true. But some forms of hypothesis tests can permit this task to be performed without any assumed probability model.[3] The most familiar of these uses is the permutation or randomization test, but nonstochastic hypothesis testing is applicable more generally. Given the data and conditions of the NTC, this is the mode that we prefer for the second of the three tasks comprising our hypothesis "tests."

Causation

Attribution of "big" observed differences to a cause or causes is not usually considered part of formal hypothesis testing, although it is implicit in and an indispensable part of the interpretation of any such test. For a given observed difference, Arroyo analysts will have their own notions of the cause of the observed difference; the purpose here is to present the hypothesis and the tests (i.e., the data evidence) to subject matter authorities and have them go through the exercise of attributing the difference to a cause or causes. This may (and, we hope, will) lead to alternative explanations of the observed differences. In some cases, these alternative explanations can be "tested" as above with easily accessible data, but in other cases, new data will need to be collected. In either case, this process of attributing the observed difference to a cause or causes will illuminate the real issues in the area under study better than simply stopping with the Arroyo analysts' formal hypothesis test.

These methods call for one more refinement when applied to policy analysis. Statistical analysis ordinarily divides data between *dependent* variables which are the outcome of the system under study--

---

[3]See, for example, D. A. Freedman and D. A. Lane, "A Non-Stochastic Interpretation of Reported Significance Levels," *Journal of Business and Economic Statistics*, Vol. 1, 1983, pp. 292-298. Other references appear in the Bibliography.

in this case, the general class of battle outcomes, which measure combat capabilities; and *independent* variables which affect the outcomes, e.g., effective use of artillery. For policy analysis, however, the independent variables must be split between *policy variables*, which can be deliberately changed in order to improve outcomes; and nonpolicy variables, which the Army cannot change. A policy variable would be the location of the battalion Fire Support Officer relative to the Commander; a nonpolicy variable would be the relative effectiveness of artillery in defensive and offensive battles. If artillery is more effective on the defense, it is important to include that fact in the system as a factor to be allowed for in moving toward policy recommendations for artillery, but among those policy recommendations will not be one that battalions avoid going on the offensive.

THE NTC DATA ANALYSIS PLAN

The policy analysis concepts discussed above translate into four operational steps being used by the Arroyo Center project to develop a methodology for derivation of NTC lessons for the Army:

- Choice of issues
- Formulation of hypotheses
- Selection and collection of data
- Use of the data to test the hypotheses

Choice of Issues

For two reasons, an NTC analytical methodology will be developed by analyzing specific Army issues in the NTC setting, rather than creating the methodology and then applying it to the issues. First, specificity works toward both objectives of the Arroyo Center NTC program, providing immediate assistance to the Army as well as developing methods for continuing assistance. And second, NTC is a innovative system; although the general methods of policy analysis and data analysis discussed above do apply, specific applications to NTC must be learned from concrete examples.

The first specific issue examined at NTC was ground-to-ground fratricide by both direct and indirect fire. Fratricide was chosen because of synergy with another Arroyo project, a study of how to reduce fratricide of all types. The findings on ground-to-ground fratricide will be the subject of a subsequent Rand Note and will also be incorporated in the overall fratricide project.

One such problem can frequently lead to others. In the course of examining indirect fire fratricide, for example, it was noted that NTC artillery had an unexpectedly low accuracy rate; whether this is real or an artifact of the methods of simulating indirect fire, and if it is real, why, are questions to be investigated.

The major impact of the initial Arroyo Center NTC effort has been the experience of working with NTC data of all the types described above, and improving understanding of data limitations as well as data-handling methods. The initial study had major limitations as a prototype of NTC project methodology, however. Fratricide (fortunately) happens only infrequently, at NTC as in real battle. For that reason, the types of data categorized above as being gathered "non-systematically" but "potentially the most valuable" were not readily available for this set of issues, and the methods used were significantly narrower than those being developed for continuing use at NTC. And in the case of the artillery, the fact that the simulation of indirect fire is substantially more artificial than that for direct fire (MILES is a line-of-sight system not usable for indirect fire, and effects must thus be estimated subjectively) places some doubt on the meaning of the NTC results.

To test a full mix of data-utilization methods, therefore, the next prototype issue chosen will be one better suited to employ the range of data actually or potentially available; and will be in an area of activity where it is believed that the simulation at NTC reproduces more closely the conditions of actual battle. Possibilities include reconnaissance/counter-reconnaissance or other issues related to Command/Control/Communications/Intelligence.

Hypothesis Formulation

Hypothesis formulation begins with a "description of the situation", as discussed above. Description is the first of four operational steps of NTC hypothesis formulation to be carried out by a group of people assembled (literally or by communications net) for the purpose. The group might include only Arroyo analysts or it might also bring in NTC or other military personnel. If it does not include the latter, the Arroyo analysts should have carried out preliminary groundwork by systematic gathering of ideas at NTC. The four steps are:

a. Describing the issue and situation operationally, in terms of concrete observable events, as discussed on pages 12-13.
b. Dividing the data measures of the events described in the first step--among dependent variables, policy variables, and other independent variables.
c. In terms of these variables, formulating hypotheses or, equivalently, choosing ways to classify or display the data about battles, rotations, or other relevant units of experience. This step may involve appropriate statistical methods like clustering or CART.
d. Repeating step (c) as often as appears useful, possibly by bringing new players--perhaps from NTC--into the game.

The end result will be a classification or other description of the data that should suggest policy hypotheses.

Take, for example, the problem of reconnaissance on offense. It is generally believed throughout the Army that a successful offense is not possible without successful reconnaissance, and that when good intelligence, including reconnaissance, is available, offensive battles are usually successful.

Once the hypothesis-generating group is assembled, the object would be to operationalize (i.e., describe in terms of concrete events that can be observed and, in some cases, measured) both the degree of success of reconnaissance and the success or failure of the offense. This does not mean that everything must be quantified. Rather, it is necessary to

find identifiable, concrete, existing phenomena that capture as much as possible of what is important about reconnaissance and battle success.

For reconnaissance, this might include:

- Whether the scouts penetrated the objective area undetected.
- Whether the scouts located enough vehicles, personnel, and obstacles to permit the battalion intelligence officer (S2) to accurately display the enemy dispositions and the directional orientation of these positions.
- Whether the scouts breached the obstacles, marked the breaches, and reported the location and type of marking of the breaches (if they had been tasked to do so).

Ideally, such a breakdown would be at once comprehensive, easy to reduce to a set of objectively answerable questions, and easy to collect.

Operationalizing battle success is more difficult and contentious. Kill ratios and attrition measures are insufficient, both because they are too narrow to reflect a sensible assessment of battle outcome, and because battles at NTC are allowed, for pedagogical reasons, to go on past the point at which one of the sides would have disengaged were the battle real. The measure of success must reflect the mission. One possible measure for offensive missions, for example, would be to declare the attack a success if the objective was physically occupied (regardless of remaining combat power on either side), and a failure otherwise. Alternatively, the offensive mission could be declared a success only if the objective were physically occupied and the occupying unit retained some given measure of combat power.

These measures are listed for illustrative purposes only; in any actual use of this methodology, the assembled subject matter experts would judge which measures to use. For example, they might not want to use overall battle outcome, but some intermediate measure such as the ability of the offensive unit to execute an adaptive feature of its battle plan, perhaps selecting a path of approach, while on the move.

To continue this reconnaissance example, the policy variables include the tasks assigned to the scouts, the vehicles and other assets assigned to the scouts, and the technology available to the units (e.g., night vision equipment). Obviously, scout training is implicit here, but if scouts are all trained in roughly the same way, training cannot be introduced explicitly into the analysis. The dependent variables are the measures of success of the battle outcome. The nonpolicy independent variables include such obvious things as weather, visibility, and terrain features, but also items like the type of attack (hasty, deliberate) and the overall size of the attacking force, which may be depleted substantially due to attrition or maintenance difficulties.

Once these lists of variables are in hand, the next step is to set forth concrete hypotheses in terms of "do this" options. Two such hypotheses might take the following illustrative forms:

1. Because scouting makes little contribution to the successful execution of a deliberate attack, *men and assets devoted to scouting for such attacks should be reduced.*
2. Because reporting the location of obstacles is of little use if the obstacles remain in place, *scouts should breach the obstacles, and maintain and clearly mark the breaches.*

These hypotheses will in turn suggest classifications of battles:

1. By type of attack: hasty, deliberate, or otherwise.
2. By success or failure at locating and reporting obstacles, and success or failure at breaching them, and maintaining and marking the breaches.

For the first hypothesis, the battles would be further classified according to success or failure of the scouting and success or failure of the offense. If success of scouting and successful battles were related for the hasty attack but not for the deliberate attack, this would be evidence in favor of the hypothesis. For the second

hypothesis, the battles would be further classified according to success or failure of the battle. If battle success were related to successful location and reporting of obstacles for those battles where the obstacles were breached and marked, but not for those battles where the obstacles were not breached and marked, this would be evidence in favor of the hypothesis.

Data Selection and Collection

The seventeen types of actual or potential NTC data classified above according to the means by which they are gathered, recorded, and retained must be categorized differently for their use in lessons-learned analyses. One crucial distinction is between:

- Data which are recorded and are *tabulated* by NTC instrumentation or other means (e.g., rounds fired, hits, vehicle kills) or are relatively easily tabulatable (special summaries of artillery logs, for example). These are referred to as *T-Data*.
- Data which are observed and recorded but would take a substantial effort to reduce to tabulatable form. Video tapes, OC notes, and After-Action Reviews are examples here. These are called *O-Data*.

Two other data distinctions are important for analytical purposes. The first distinction is between *Descriptive* data, stemming directly from the battle, which can be used to describe specific aspects; and *Analytic* data, the records made by individuals as part of their analytical explanations to themselves or others of what happened and why. The second distinction is between data deliberately *Planned* to assist in the specific analysis at hand, and *Unplanned* data, gathered routinely or for other purposes but which may be used in the analysis. This two by two by two (T-Data/O-Data; Descriptive/Analytic, Planned/Unplanned) classification results in eight cross-categories. Table 1 divides the seventeen data types among these categories; certain types appear in more than one place, depending on how they are to be gathered and utilized.

For the use of the data to test hypotheses, discussed below, the crucial distinction is between T-Data and O-Data. For data collection, however, the other categorizations also assume importance. Obviously, plans must be laid to gather planned data. Special-purpose interviews, for example, which appear in all four of the planned-data categories--depending on whether the interviews of OCs or other knowledgeable officers are intended to gather numbers or more general observations (T-Data or O-Data), and whether they are intended to elicit descriptions or viewpoints (Descriptive or Analytical)--must be carefully laid out

Table 1

DATA TYPES BY ANALYTICAL CATEGORIES

Data	T-Data	O-Data
Descriptive		
Planned	13. Special-purpose interviews	13. Special-purpose interviews
Unplanned	2. Real-time digitals 3. Computer summaries 4. Unprocessed data 11. OC kill and related records 12. Artillery logs 15. Pre- and post-NTC data	1. Graphics 5. Communications tapes 6. Battle video tapes 8. Scenarios 9. Operations orders
Analytic		
Planned	13. Special-purpose interviews 14. Analyst observations	13. Special-purpose interviews 14. Analyst observations 17. Officer observations
Unplanned	None	7. AAR videos 10. Take-home packages 16. OC Notes

both for content and for non-interfering approaches to interviewees. And for unplanned data, arrangements must be made for retention.

Perhaps the most important category of data, however, contains the analyst observations which appear in the *Analytic/Planned* category for both T-Data and O-Data. These are the field observations of the Arroyo analysts themselves and of the CATA officers working with the Arroyo analysts. Numbers (T-Data) alone can be misleading and it is important to supplement them with observations of trained observers about "what really happened." In the fratricide analysis, for example, the numbers and the graphic observations indicate what is likely to have happened--e.g., a fratricidal hit or near miss at very short range probably indicates a round fired accidentally--but the information would have been more definitive had a trained observer been on the spot to find out "why" as well as "what."

For this type of personally gathered data, substantial emphasis must be put on training data collectors and planning for data collection. Arroyo analysts must be well trained in understanding what goes on in the Army as a whole and at the NTC; all data collectors, Army and civilian alike, must understand both the characteristics of the NTC simulation and the requirements of systematic data gathering. Trained observers of either type, however, are a scarce and expensive resource; their utilization must be carefully planned to gather the data most needed to illuminate the specific hypotheses under consideration, rather than merely "sending them out to the field" to make general observations about what is going on.

Different issues call for different mixes of data types. The fratricide analysis omitted analyst observations for a very good reason: fratricide is so infrequent that it would have been wasteful in the extreme to have observers standing around waiting for it to occur. Rather, the fratricide analysis used only *Descriptive/Unplanned* data, both T-Data (fratricide summaries compiled for each engagement by the computer system) and O-Data (review of the graphics for the fratricidal incidents to try to figure out the circumstances). Had the fratricide analysis taken place after data-retention routines now being planned were in operation, such *Analytical/Unplanned* data as OC notes would also have been extremely helpful.

This analysis has provided some immediate benefits to the Army in terms of lessons learned. Nonetheless, it was much too narrow to test the full range of data collection and data utilization efforts needed to develop a methodology making optimal use of data gathered and gatherable at NTC. The next problem to be examined will be chosen with such methodological development a prime consideration.

Hypothesis Testing

As has been noted, this step is far from the familiar testing of hypotheses carried out by natural scientists or even practitioners of formal statistical analysis. Rather, what will be done is to examine the hypotheses in the light of the data gathered, to search for what we have called "big" differences, and to attribute those differences to causes. This kind of testing takes place in three stages, as noted earlier.

In the first stage, Arroyo analysts will describe (i.e., reduce) the data in some relevant manner and perform formal (preferably nonstochastic) hypothesis tests on interesting observed differences. For instance, consider the second hypothesis from the scouting example--that reporting the location of obstacles is of little use if the obstacles are not breached and the breaches are not marked and maintained. The following classification of offensive missions (i.e., description of the data) is suggested by the hypothesis: classify each offensive mission according to the success or failure (i) of the scouts at locating and reporting obstacles, (ii) of the scouts at breaching the obstacles, and marking and maintaining the breaches, and (iii) of the offensive mission itself.

In the second stage, the formal hypothesis test, Arroyo analysts could compute a particular measure of association (many are available) between classifications (i) and (iii), for missions for which (ii) was successful, and again for missions for which (ii) was not successful. The two observed measures of association for these two collections of offensive missions will differ--if the hypothesis is true, (i) and (iii) will be associated for missions in which (ii) is successful but not otherwise--and a permutation test could be used to judge whether the

observed difference between the measures of association for the two collections of offensive missions was "big" ("significant" in more traditional jargon).

Whether an observed difference is judged "big" or not, the analysts' formulation and the test will contain an implicit message--that breaching, marking, and maintaining obstacles is important (or not, depending on how the permutation test turns out). In the third stage, these hypotheses and the implicit message of the Arroyo analysts' tests will themselves be tested against the experience and intuitions of NTC and other military personnel. In effect, these subject matter specialists will be repeating the problem formulation stage to make their own judgements of the causes to which observed differences (or the absence of differences) should be attributed. The process is one which in academic circles would be called a seminar but in military terminology might be considered a structured debriefing. Unlike an AAR, which in some ways it resembles, such a debriefing would be devoted to presentation and discussion of the specific hypotheses arrived at in the first stage. Participants would be encouraged to bring in their own insights and intuitions and support them with whatever kinds of T-Data or O-Data seem appropriate. Replay of engagement graphics might be part of the process.

For each hypothesis under discussion, the debriefing results would range between two poles:

- Hypotheses for which the outcome of the analysis tends to *agree* with the intuitions of the participants. Given all the imperfections of the process, this would be the closest the project could come to actually *confirming* a hypothesis. In the reconnaissance case, for example, the hypothesis that good reconnaissance and battle success are very highly correlated might be readily accepted by practically all officers.
- Hypotheses for which the outcome of the analysis tends to *disagree* with the intuitions of the participants. In some ways, this is the more fruitful end of the spectrum. It is a classic clash between formal analysis and experienced

judgment. The proper analytical bias here favors intuition; for all of our statistical sophistication, the experienced human mind carries out a powerful analytical process comparable to and frequently better than more formal processes. The crucial point, however, is not whether judgment is to be preferred over more formal analysis; it is that the debriefing session should be used to understand why the analysis and the informed judgments disagree, and by this process to understand what happened at NTC and what might happen in real battle. More frequent and more important insights can come from this process than from analytical confirmation of what might be obvious in any case. Should one of the reconnaissance hypotheses mentioned above, for example--that scouting makes little contribution to the successful execution of a deliberate attack, so men and assets devoted to scouting for such attacks should be reduced--appear to be confirmed by data analysis, it is likely to raise a lot of judgmental opposition, but the subsequent discussion might well throw some real light on the risks which should be taken for successful reconnaissance under varying circumstances.

Finally, then, the Arroyo team will put together what they believe and why they believe it, using the data and discussions to formulate the basic "line of argument." By this process, we believe that the NTC data will add a dimension hitherto missing from peacetime analyses of Army doctrine, and thereby help the Army improve its combat capabilities. The proof, however, is not in this planning description; it will be in the utilization and refinement of this methodology as it is applied to specific Army needs.

BIBLIOGRAPHY

Breiman, Leo, Jerome H. Friedman, Richard A. Olshen, and Charles J. Stone, *Classification and Regression Trees* (Belmont, CA: Wadsworth, 1984).

Cox, D. R., "The Role of Significance Tests," *Scandinavian Journal of Statistics*, Vol. 4, 1977, pp. 49-70.

Finch, P. D., "Description and Methodology in the Practice of Statistics," *Biometrika*, Vol. 66, 1979, pp. 195-208.

Freedman, D. A., and D. A. Lane, "A Non-Stochastic Interpretation of Reported Significance Levels," *Journal of Business and Economic Statistics,*, Vol. 1, 1983, pp. 292-298.

Freedman, D. A. and D. A. Lane, "Significance Testing in a Non-Stochastic Setting", *Festschrift for Erich Lehmann* (Belmont, CA: Wadsworth, 1983), pp. 185-208.

Jones, Donald R. and Kim Henrickson, *Development of a Leader Lessons Learned Methodology*, Allen Corporation of America, Report submitted to the U.S. Army Research Institute, October 1981.

Mallows, Colin L., "Data Description," in G.E.P. Box, T. Leonard, and C.-F. Wu (eds.), *Scientific Inference, Data Analysis, and Robustness* (New York: Academic Press, 1983).

Rouanet, H., J.-M Bernard, and B. Lecoutre, "Non-Probabilistic Statistical Inference: A Set-Theoretic Approach," *American Statistician*, Vol. 40, No. 1, February 1986, pp. 60-65.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: N-2461-A	2. GOVT ACCESSION NO.
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle): Utilizing the Data from the Army's National Training Center: Analytical Plan	5. TYPE OF REPORT & PERIOD COVERED: Interim
	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s): Robert A. Levine, James S. Hodges, Martin Goldsmith	8. CONTRACT OR GRANT NUMBER(s): MDA903-86-C-0059
9. PERFORMING ORGANIZATION NAME AND ADDRESS: The Rand Corporation, 1700 Main Street, Santa Monica CA 90406	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS: Office of the Deputy Chief of Staff, Headquarters United States Army, Washington DC 20310	12. REPORT DATE: June 1986
	13. NUMBER OF PAGES: 27
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)	15. SECURITY CLASS. (of this report): Unclassified
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for Public Release: Distribution Unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

No Restrictions

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Army Training	Transfer of Training
Land Combat	Military Doctrine
Army Operations	Data Acquisition

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

See reverse side

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

This Note describes the Army's operations at its National Training Center (NTC) at Fort Irwin, California, and ways in which the data collected during those operations can be used to derive lessons about Army doctrine, training, and weapon systems. The discussion of operations at the NTC includes descriptions of the training conducted there, the facilities for training and data collection, and the types of data actually or potentially available. As a laboratory for deriving lessons, the NTC has unique advantages and disadvantages. These have implications for using the NTC experience to formulate and test hypotheses. In particular, the authors emphasize the importance of testing the results of formal analytic procedures against the experience and intuition of NTC and other military personnel. Keywords: Army training, Land combat, Transfer of training, Data acquisition