

DTIC F-100Y

Technical Report 825

Review of Research and Methodologies Relevant to Army Command and Control Performance Measurement

Lloyd M. Crumley
U.S. Army Research Institute

January 1989

DTIC
ELECTE
AUG 14 1989
S D CS D



United States Army Research Institute
for the Behavioral and Social Sciences

Approved for public release; distribution is unlimited.

89

8

14

002

AD-A211 247

U.S. ARMY RESEARCH INSTITUTE FOR THE BEHAVIORAL AND SOCIAL SCIENCES

A Field Operating Agency Under the Jurisdiction
of the Deputy Chief of Staff for Personnel

EDGAR M. JOHNSON
Technical Director

JON W. BLADES
COL, IN
Commanding

Technical review by

Thomas O. Jacobs
William D. Sprenger, Shippensburg Univ.
Shippensburg, PA

Accession For	
NTIS CRA&I	☒
DTIC TAB	☐
Unannounced	☐
Justification	
By	
Distribution	



A-1

NOTICES

DISTRIBUTION: Primary distribution of this report has been made by ARI. Please address correspondence concerning distribution of reports to: U.S. Army Research Institute for the Behavioral and Social Sciences, ATTN: PERI-POX, 5001 Eisenhower Ave., Alexandria, Virginia 22333-5600

FINAL DISPOSITION: This report may be destroyed when it is no longer needed. Please do not return it to the U.S. Army Research Institute for the Behavioral and Social Sciences.

NOTE: The findings in this report are not to be construed as an official Department of the Army position, unless so designated by other authorized documents.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1a. REPORT SECURITY CLASSIFICATION Unclassified			1b. RESTRICTIVE MARKINGS --	
2a. SECURITY CLASSIFICATION AUTHORITY --			3. DISTRIBUTION / AVAILABILITY OF REPORT Approved for public release; distribution unlimited.	
2b. DECLASSIFICATION / DOWNGRADING SCHEDULE --				
4. PERFORMING ORGANIZATION REPORT NUMBER(S) --			5. MONITORING ORGANIZATION REPORT NUMBER(S) ARI Technical Report 825	
6a. NAME OF PERFORMING ORGANIZATION ARI Field Unit at Fort Leavenworth, Kansas	6b. OFFICE SYMBOL (If applicable) PERI-SL	7a. NAME OF MONITORING ORGANIZATION U.S. Army Research Institute for the Behavioral and Social Sciences		
6c. ADDRESS (City, State, and ZIP Code) P.O. Box 3407 Ft. Leavenworth, KS 66027-0347		7b. ADDRESS (City, State, and ZIP Code) 5001 Eisenhower Ave. Alexandria, VA 22333-5600		
8a. NAME OF FUNDING / SPONSORING ORGANIZATION --	8b. OFFICE SYMBOL (If applicable) --	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER --		
8c. ADDRESS (City, State, and ZIP Code) --		10. SOURCE OF FUNDING NUMBERS		
		PROGRAM ELEMENT NO. 63739A	PROJECT NO. 793	TASK NO. (1307) 149
				WORK UNIT ACCESSION NO. H1
11. TITLE (Include Security Classification) Review of Research and Methodologies Relevant to Army Command and Control Performance Measurement				
12. PERSONAL AUTHOR(S) Crumley, Lloyd M.				
13a. TYPE OF REPORT Final	13b. TIME COVERED FROM 87/01 TO 88/03	14. DATE OF REPORT (Year, Month, Day) 1989, January	15. PAGE COUNT	
16. SUPPLEMENTARY NOTATION --				
17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)	
FIELD	GROUP	SUB-GROUP	Command Performance	
			Control Evaluation	
			Command post Command and control	
19. ABSTRACT (Continue on reverse if necessary and identify by block number) This report reviews research that has been conducted on division-, brigade-, and battalion-level command and control processes. Battalion-level research, which forms the largest part of the literature, is shown to have been motivated primarily by the need to evaluate battalion Army Training and Evaluation Programs (ARTEPs) and factors related to battalion battle simulation development. The literature on research and methodology development is also shown to address certain basic issues such as what constitutes command post effectiveness and how battle outcome measures relate to command group performance. Some attention is also given to problems associated with the conduct of research in this arena and the development of research, or applied, measurement processes.				
20. DISTRIBUTION / AVAILABILITY OF ABSTRACT ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS			21. ABSTRACT SECURITY CLASSIFICATION Unclassified	
22a. NAME OF RESPONSIBLE INDIVIDUAL Lloyd M. Crumley			22b. TELEPHONE (Include Area Code) (913) 684-4933	22c. OFFICE SYMBOL PERI-SL

Technical Report 825

**Review of Research and Methodologies
Relevant to Army Command and Control
Performance Measurement**

Lloyd M. Crumley
U.S. Army Research Institute

ARI Field Unit at Fort Leavenworth, Kansas
Seymour M. Halpin, Chief

Systems Research Laboratory
Robin L. Keesee, Director

U.S. Army Research Institute for the Behavioral and Social Sciences
5001 Eisenhower Avenue, Alexandria, Virginia 22333-5600

Office, Deputy Chief of Staff for Personnel
Department of the Army

January 1989

Army Project Number
2Q263739A793

Human Factors in Training
Operational Effectiveness

Approved for public release; distribution is unlimited.

FOREWORD

The U.S. Army Research Institute for the Behavioral and Social Sciences (ARI) is conducting research to develop improved methods for measuring the performance of command staffs at the corps, division, and brigade levels. A part of that effort has been devoted to a review of published material that deals with research on Army command and control staff performance. It was found as the review progressed that the literature was widely scattered. It also became apparent that there was a wide range in the quality of the research. This review presents the reviewed literature in some detail and calls attention to the strengths and weaknesses of reported research when it is appropriate. The detail and commentary should be of help to researchers and workers in the area who wish to become familiar with the field.

This effort is part of the Fort Leavenworth Field Unit's research program to develop improved command and control staff measurement techniques. The Field Unit's overall mission is to conduct research that will aid the Army in improving its command and control processes. This research was conducted as part of Research Task 1.4.9, Evaluating Command Post Performance. A Letter of Agreement entitled, "Development of Performance Measurement Methodology for Corps Division and Brigade Command Posts," which supports this task was signed by the Director of the Combined Arms Combat Development Activity (CACDA), Command, Control, Communications and Intelligence (C³I) Directorate and the Commander, US ARI on 11 July 1986. Material from this report has been the basis for presentations to MG Sullivan, Deputy Commandant of the Command and General Staff College (CGSC) (29 June 1987); COL Dacunto, Director, C³I (16 June 1987); COL Brown, Deputy Director, ADEA (22 July 1987); and LTC Norris, Chief BCTP, AAR Division (31 August 1987). Potential users who have been made aware of this review at these and other presentations have indicated that they foresee a considerable use of this product as a resource for persons working in the command and control area.



EDGAR M. JOHNSON
Technical Director

REVIEW OF RESEARCH AND METHODOLOGIES RELEVANT TO ARMY COMMAND AND CONTROL PERFORMANCE MEASUREMENT

EXECUTIVE SUMMARY

Requirement:

The literature which deals with command and control performance research and measurement methodologies is sparse, scattered, and of highly variable quality. As a result of these factors, much of what has been done by earlier researchers is either unknown to, or ignored by, later researchers. Because the U.S. Army Research Institute for the Behavioral and Social Sciences (ARI) has initiated a research project to develop improved measures of command and control performance in Army command posts at the corps, division, and brigade level, it has become imperative that a document be created and made available to interested parties, so that effective use can be made of past findings as new research is planned and conducted.

Procedure:

The procedure involved in this literature review was simple and direct. The ARI Fort Leavenworth Field Unit technical library was searched and the researchers at the unit were consulted to obtain documents and titles that were known to them but not in the library. Additional material was obtained by searches of the Command and General Staff College library and by consulting standard abstracting sources such as the National Technical Information Service (NTIS). Finally, as published reports were obtained, their lists of references were reviewed to determine if prior authors had identified documents that were not yet known to the reviewer. The literature was then read, evaluated, and discussed in the context of the presentation plan developed as the available material was evaluated.

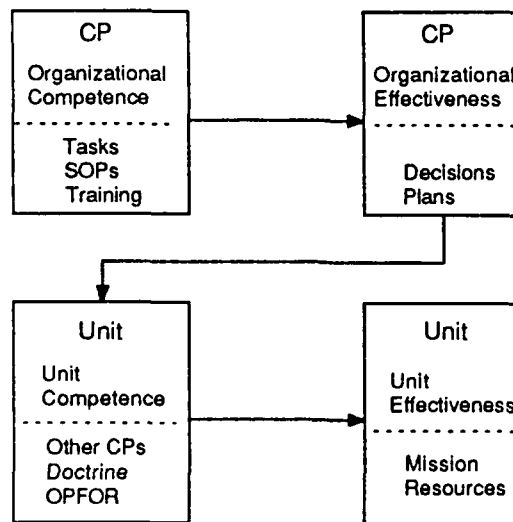
Findings:

Most of the literature that reports command and control research was done on battalion-level command groups and deals with Army Training and Evaluation Program (ARTEP) and/or battle simulation issues. Despite these limitations, research done to date has addressed issues that are important in the command and control performance measurement area. The earliest, and probably the finest, research considered the problem of what constitutes a suitable criterion of command post effectiveness. The researchers introduced the concepts of "competence" and "effectiveness." In their research organizational competence was determined by a series of measures based on the processes used to define the command and control cycle, and organizational effectiveness was

determined by scores based on the decisions arrived at by the tested command group. Results of a carefully conducted, laboratory-like data collection and analysis process demonstrated that commanders supported by better performing--more competent--command staff almost without fail made better decisions than commanders supported by less competent command staffs.

Later research, by the same and other researchers, demonstrated that the measured relationships between command group performance and command group effectiveness seriously decreased when subjective measures (such as ARTEP task ratings) replaced the most objective measures used in the original study and disappeared almost completely when battle outcome data were substituted for decision quality data.

In general the various research projects that have attempted to find a relationship between command group task performance and battle outcome data have demonstrated that the relationships are not strong. Indeed, sophisticated statistical techniques, such as factor analyses, have suggested that in the case of ARTEP-based task data the extensive correlational matrices show only that the individual task scores are at worst much biased by a halo effect caused by the controller/rater's knowledge of how the battle is going. Even at best, these data support only two factors consisting of (1) the tasks involved in planning, and (2) the tasks involved in controlling performance against the plan. Since a few studies gave results where the best--almost significant--correlations tended to show that improved command performance made for worse simulated battle performance, it seems safe to conclude that command staff competence, objectively measured, predicts the effectiveness level of the command and control process at the command post, but that neither command group competence nor command group effectiveness are adequate predictors of battle outcome. Thus, past research tends to support a command and control performance measurement model such as this one,



which considers CP effectiveness apart from unit effectiveness.

Past research clearly demonstrates that measures of the competence of a command and control organization can be validated by demonstrating that they correlate with command and control effectiveness measures that are command post specific. Such items as making the right decision, preparing orders that last, doing the doctrinally correct thing in task organizing or assigning missions. The broader question of how these command post specific things relate to battle outcome is much more complex and raises additional questions such as, "Is doctrine in this area correct?" and "What can the enemy do to counter my best-laid plans?"

Research has demonstrated that decision quality and plan quality can both be used as command post effectiveness measures. However, the same research has shown that there are difficulties associated with both approaches. The decision quality measurement approach has only been demonstrated at battalion level. To apply the technique, researchers first had to develop a large number (128) of scenario events that could be presented to a battalion headquarters. Researchers then needed to determine all the possible responses, rate each response alternative for goodness, rate each event for difficulty, and then create scenarios that incorporated the events. Successful application, therefore, required both development of the necessary material and control of the scenario. Results using this approach were very successful and provided the command and control arena's most outstanding research.

Research that directly scored plans in terms of the soundness of the decisions incorporated in them has also been successful. This technique, which has been applied to division-level command and control research, requires that there be a "school-house" solution against which each plan can be compared. Application of the methodology also requires that potential parameters of deviation--task force organization, phase lines, etc.--be identified and scales prepared for each. This method is somewhat limited in its potential for application because it requires a controlled scenario, creation of a "correct" solution, and development of scales to evaluate potential deviations on all of the parameters that might vary.

A methodology that evaluates the effectiveness of plans at any echelon by indirect measures has shown signs of being the most significant development in the field. This approach makes use of the concepts implicit in the four-step model described earlier. The method presumes that a unit commander is attending to the battle as it develops and that he will have a new plan prepared when he believes that the current plan is not effectively addressing the developing battle situation. The method assumes that plans that last longer are superior to plans that do not last as long. This leads to a primary command and control effectiveness measure--plan life. The researchers also assume that a command and control organization that, when it needs to replan, is able to create a new plan without having to "start from scratch" and, thereby, have to complete a full planning cycle, is more effective than a command group that is pretty much taken by surprise when the plan fails and must complete a full planning cycle. This approach, supported by an appropriate Adaptive-Coping-Cycle-based planning cycle model, shows considerable promise.

Attempts to directly score battalion level operation orders (OPORDs) by methods that consider specific contents, such as clarity, information content, etc., have also been reported. The research that developed and/or used one of the two existing methods has proven to be of limited value. The primary problem is the context in which the research was conducted. The reported research attempts to determine how OPORD quality relates to battle outcome as determined in a simulation-based command post exercise (CPX) where only four battalion staffs were tested. Thus, the research falls prey to the problems associated with attempting to find relationships, likely to be weak at best, with inadequate data. It is probably safe to conclude that the two OPORD evaluation scales that have been developed would be better used in situations where relationships to task performance within the command staff were being investigated.

Much of the battalion command and control research and some brigade research was motivated by the needs of persons developing simulators and the ARTEP system that was concurrently evolving. Researchers, always alert for an opportunity to collect data, were also able to conduct meaningful projects while investigating the ARTEP/simulator arena. A variety of studies have addressed the potential of the evolving battalion simulations to provide a vehicle for conducting CPXs and ARTEPs. Included were efforts that assessed comparative effectiveness and cost in evaluating the Computer-Assisted Map Maneuver System (CAMMS) versus a conventional CPX. The research showed that simulation-based exercises were cost effective, provided better training, and were better received as training events than ordinary CPXs.

Other research demonstrated that battle simulations, such as CAMMS, could be used to conduct ARTEPs at both brigade and battalion level. The research showed that ARTEPs could be conducted on CAMMS but that not all subtasks could be evaluated. The problem of non-scorable subtasks was sometimes the result of the tasks themselves and, less often, of a "non-playable on CAMMS" problem. Other projects have provided data that support similar conclusions regarding a competing battalion-level simulation, CATTS (combined arms tactical training simulator). Still later command and control researchers at ARI provided data, obtained on CATTS, to demonstrate that the evolving battalion combat simulation known as ARTBASS (Army Training Battle Simulation System) would be a cost-effective and satisfactory training system.

As these projects were conducted and reported, it became clear that battle simulations were valuable training situations and that, although there were some problems, ARTEPs could be conducted using simulations rather than traditional CPXs. Other analyses conducted on data from these and other projects of the same type did provide some disquieting results. Most notable were continuing indications that better ARTEP task performance did not always assure better battle outcome scores. A second problem was that the ARTEP raters did not appear to be good scorers of individual ARTEP subtasks. Raters, who were usually also controllers, noted that they were too busy to give the rating process the attention it required. Attempts to provide easy-to-score scales, better definition of ratings, and other aids were never fully successful. In the end the individual subtask ratings were much too correlated with one another and clearly contaminated by the fact that the controllers tended to be unduly influenced by the state of the battle being fought.

The research essentially demonstrated that conducting command and control research as an add-on to ARTEPs performed in battle simulations was possible but that these situations did not provide a good research environment. It also became clear as time passed that sound statistical analyses were needed to capture the meaning of the resulting data and that attempts to relate ARTEP tasks to battle outcomes were tenuous at best. Conclusions of this nature supported the concept, articulated by a series of ARI-funded research efforts, that good command and control research requires a dedicated facility where researchers could control command group activities and the data collection process.

Despite the problems associated with piggy-backed experiments, some very interesting research has been conducted in ARTEP/simulation situations in which researchers were able to develop hypotheses that could be addressed by data naturally flowing from the events or create relationships that enabled them to superimpose a research project on the training event. Examples are research on information flow and research that modified and expanded the earlier organizational process/effectiveness model and developed evidence to support earlier findings that increased competence of the command group in performing its tasks leads to improved effectiveness in selecting better courses of action. These are typical of studies demonstrating that with good cooperation a research event could be run in conjunction with command and control evaluations intended primarily for some other purpose. Later research by personnel at the ARI Field Unit at Fort Leavenworth also showed that data from only partially controlled events could be used in research to test command-and-control-related hypotheses.

It is probably safe to conclude that the numerous reports that tried to "do science" with data taken from beyond the behavioral science laboratory have shown several things. One is that great care needs to be taken to secure the willing cooperation of trainers or other nonresearchers who have goals that may be affected by a superimposed research goal. A second conclusion evolves out of the difficult nature of the command and control research area and the additional difficulties imposed when attempts are made to inject such research into some other event. Certainly the body of research demonstrates the need to apply sophisticated statistical techniques (e.g., factor analysis rather than mere correlations) and demonstrates that inept experimental design and naive data analyses can lead either to failed experiments or improper conclusions. Overall it seems that data from training or field events have the potential for making major contributions to the knowledge of command and control. But, at the same time, it appears that the availability of such data places great demands on researchers to assure that sound scientific methods are applied to such data to keep it from becoming a source of misinformation.

Utilization of Findings:

Research and development in command and control is expensive, complex, and difficult. The command and control area also suffers because researchers and other professionals are often thrust into responsible positions with

little opportunity to become aware of the existent body of research. This report will make it possible for interested parties to survey the field and become aware of the extent and limitations of the research foundations of the command and control performance measurement process in a relatively short time. It will be of value as background material for all persons involved with the command and control performance measurement process.

REVIEW OF RESEARCH AND METHODOLOGIES RELEVANT TO ARMY COMMAND AND CONTROL PERFORMANCE MEASUREMENT

CONTENTS

	Page
INTRODUCTION	1
BATTALION COMMAND AND CONTROL RESEARCH	6
In the Beginning	6
Research Hiatus	20
Olmstead Revisited	24
CATTs Based Research	32
Another Benchmark Study?	40
Who Knew What, When?	43
Simulators and ARTEPS	49
BRIGADE AND DIVISION COMMAND AND CONTROL RESEARCH	68
Brigade Level Research	68
Division Level Research	77
OTHER COMMAND AND CONTROL RESEARCH	82
Headquarters Effectiveness Assessment Tool (HEAT)	82
HEAT Applications	85
HEAT Evaluation	99
SUMMARY	100
Command Post Effectiveness. What to Measure?	100
Research on Battle Outcomes as Measures	103
Product Quality Based Effectiveness Measures	105
ARTEP Related	106
REFERENCES	109
APPENDIX A. PROCESS PERFORMANCE ASSESSMENT	A-1
B. PROCESS DEFINITIONS AND CRITERIA	B-1
C. SUMMARY PROCESS ASSESSMENT FORM	C-1
D. CONTROLLER RATING FORM	D-1

CONTENTS (Continued)

	Page
APPENDIX E. LIST OF TERMINAL PERFORMANCE OBJECTIVES (TPO) AND ENABLING OBJECTIVES (EC) FOR CATTs	E-1
F. FLOW CHARTS OF THE CATEGORIES OF ACTIVITIES FOR EACH PROCESS	F-1
G. INFORMATION ITEMS EVALUATED	G-1
H. CRITICAL BATTALION COMMAND GROUP TASKS, AND NUMBERS OF SUBTASKS, DESCRIBED IN ARTEP 71-2	H-1
I. PERFORMANCE STANDARDS FROM BATTALION ARTEP	I-1
J. BRIGADE ARTEP SUBTASKS	J-1
K. TERMINAL LEARNING OBJECTIVES AND LEARNING OBJECTIVES FOR BRIGADE COMMAND GROUPS	K-1
L. UNIT RATING FORM--CAMMS EXERCISE	L-1
M. SCORING OF ASSISTANT G3 PLANS SOLUTION	M-1
N. HEAT AS A METHODOLOGY EVALUATION	N-1

LIST OF TABLES

Table 1. Correlations between major variables and sub-variables	14
2. Correlations between effectiveness, competence, and process measures	16
3. Summary of multiple correlation between processes and effectiveness	17
4. Correlations between organizational process scores	28
5. Correlations between combat outcome scores and organizational process scores	29
6. Comparison of combat outcome scores between more and less process effective battalions	31
7. Overall planning scores by battalion and day of exercise	36

CONTENTS (Continued)

Page

LIST OF TABLES (Continued)

Table 8.	Relative loss exchange ratio scores	38
9.	Model and related performance standards for the military decision-making process	39
10.	Statistical correlations between RER and battalion data	42
11.	Correlations of qualitative measures	43
12.	Number of required items each person should have received and persons to whom receivers should have sent information items	45
13.	Intragroup communication percentage of available required items recalled by each receiver averaged over 13 battalion command groups	47
14.	Mean information flow scores for exercise days 2 and 4	48
15.	Effect of training with feedback on the mean percent of required information received/ remembered	49
16.	Statistically significant correlations with ARTEP ratings for the five battlefield outcome measures having the greatest number of significant correlations	52
17.	Inter-item correlations of performance and difficulty items	63
18.	Summary of statistical significance level of five main variables for the four simulation outcome measures	65
19.	Correlation of process dimensions with brigade combat effectiveness and command group effectiveness	74
20.	Process performance of brigades with most and least effective command groups	75

CONTENTS (Continued)

Page

LIST OF TABLES (Continued)

Table 21.	Correlations of supervision and control dimensions with brigade and command group effectiveness	76
22.	Supervision and control in brigades with most and least effective command groups	76
23.	Product-moment correlation between nine predictors and the criterion	80
24.	HEAT measures used in experiment comparing geographic and functional organizations	89
25.	Mean values of selected HEAT scores	93
26.	Mann-Whitney U Test applied to selected DSI 1986 data	95
27.	Description of measures defined for MCS evaluations	97

LIST OF FIGURES

Figure 1.	Schematic of the Olmstead et al. approach	11
2.	Mean probe competence scores for high and low effectiveness groups under environmental pressure	18
3.	Development of Carmonette	21
4.	Family tree of Army C2 Training systems	22
5.	Model of the production of a combat order	34
6.	Flow of information from 0900 to 1700 hours	45
7.	Flow and remembering of information during a CATTS cycle	46
8.	Flow diagram of the battalion command group ARTEP	55
9.	Summary of subtasks evaluated in CATTS	56

CONTENTS (Continued)

	Page
LIST OF FIGURES (Continued)	
Figure 10. Sample of data collection sheets	58
11. Examples of controller rating sheets, item data and analysis results	59
12. Possible status of plan to situation	83
13. Headquarters structures tested at NPGS	86
14. Communications network used in role and specialization experiment	91

REVIEW OF RESEARCH AND METHODOLOGIES RELEVANT TO ARMY COMMAND AND CONTROL PERFORMANCE MEASUREMENT

INTRODUCTION

The purpose of this report is to provide a review of research that has been done on the measurement of Army command and control performance. Documents reviewed have included government agency reports and the reports of contractors. Both types of documents include material which has not received broad circulation. Documents which are not generally available, but which have been included in this review, include agency informal reports, such as Army Research Institute Working Papers, and material from contractors such as interim contractor reports. These latter types of reports are often not easily available and even in cases where they have been deposited into the Defense Technical Information Center (DTIC) system the research portion is often buried in a much broader volume which makes it unlikely a researcher would realize that command and control research was reported therein. The open literature was also searched but command and control of Army units is not a popular subject in the more public research reporting vehicles.

It was the intent of the reviewer to write a report that discussed past command and control research in sufficient detail that an interested reader would see it as an introduction into the area rather than as just an identification of the documents he or she might want to obtain and read. In order to accomplish this goal two decisions were made which have affected the nature and content of the review. The first decision was that the interested reader, by fiat, was considered to be a trained potential researcher preparing to do research in some command and control area, or a trained professional, such as an Army officer or engineer, with a significant need to be aware of the extent and nature of command and control research. Therefore, to aid the interested user in identifying research of interest, the review contains considerable detail and provides judgments concerning the quality of the research. The second decision was to limit the review as much as possible to reports which discuss data actually obtained on persons performing Army command and control processes.

The review excludes reports in a variety of areas which a broader definition of command and control research might have made relevant. Readers interested in a less constrained evaluation of the field can find reviews in a variety of areas. For example, Cooper, Shiflett, Korotkin, and Fleishman (1984) reviewed techniques for assessing team performance in command and control, Kaplan (1987) reviewed lessons learned in research on command group training, Solick and Lussier (1988) discussed what command and control research had shown needed to be considered in designing battle simulations for command group training, Sulzen (1986) has provided an annotated bibliography on tactical engagement simulations, and Garlinger and Fallesen (1988) have reviewed command group training measurement methods. Some of these reviews considered,

from a different perspective, many of the same reports discussed in this review, e.g., Kaplan (1987). Others, such as Cooper et al., go much further afield and, indeed, only two reports discussed in this review appear in a 13.5 page list of references found in Cooper's review.

This review contains four major sections. In the first, battalion level command and control research is addressed. Most of the work done to date has been done at this echelon. A second section addresses the brigade and division level literature. Only a few research efforts have been conducted at these levels and none have been reported at corps. A third section discusses applications of a specific technique, the Headquarters Effectiveness Assessment Tool (HEAT) to naval command post exercises and in small research projects. This HEAT based material is within the scope of the review because the method has been applied to Army division CPXs and more applications are planned. The final section discusses selected problems, research interests and concepts which have been identified as the review proceeded.

It might appear, based on the sharp highlighting in this report of the shortcomings of some research efforts, that little has been learned about command and control performance during the past two decades. This is not the case. There have in fact been some findings, and some products, that make a considerable contribution to research and development in the command and control performance measurement arena. Indeed some positive conclusions can be drawn from what is both a sparse and sometimes marginal literature. The major findings of the research discussed in the following review appear to be that:

1. Defining command and control activity in terms of an adaptive coping cycle model and using command decision correctness as the command post effectiveness criterion provides a sound basis for conceptualizing the relationship between a command staff's competence as an organization and its effectiveness in providing the commander the inputs he needs to make better decisions.

Olmstead, Christensen and Lackey, (1973)

2. The correlation between organizational competence and organizational effectiveness is considerably lower (.67 vs. .93) when subjective performance measures and battle outcome results, which are unit not command post specific measures, are used to determine the relationship between staff performance and command effectiveness.

Olmstead, Elder and Forsyth (1978)

3. Other research demonstrates that the adaptive coping cycle model in conjunction with the organizational competence/organizational effectiveness paradigm can be applied to various types of organizations and that different versions of adaptive decision cycle models can be applied equally well to the command and control evaluation process.

Olmstead and Christensen (1973)

Olmstead, Christensen, Salter and Lackey (1975)

Olmstead, Baranick and Elder (1978c)
Defense Systems, Inc. (1986e)

4. Data/information flow studies in battalion command posts suggest that staff personnel often do not fully obtain, transfer or, indeed, remember needed information. These studies suggest that improving the flow of information in command posts needs to be a prime consideration in preparing SOPs and conducting training exercises.

Kaplan (1980)
Thomas, Kaplan and Barber (1984)

5. Studies at both battalion and brigade level command posts show that simulator based exercises, e.g., Combined Arms Tactical Training Simulator (CATTS), Army Training Battle Simulation System (ARTBASS), provide more effective, less expensive and better received training than non-simulation based CPXs.

Barber and Kaplan (1979)
Kaplan and Barber (1979a)
Olmstead, Baranick and Elder (1978b)
Barber and Solick (1980)

6. Care needs to be taken in defining ARTEP tasks. Researchers have found that approximately 25% of ARTEP subtasks either could not be evaluated or could not be played during ARTEPs. Researchers have generally had to omit and revise subtasks because of this problem.

Barber, McGrew, Stewart and Andrews (1979)
Kaplan and Barber (1979b)

7. Attempts to correlate ARTEP task and subtask scores to subjectively obtained overall command group performance judgements have been less than satisfactory. Researchers have noted that ARTEP scorers, who inevitably are also simulation controllers, indicate that they are often too busy to do a good evaluation job. There is also evidence of a strong halo effect since controllers who must score subtask performance appear to be influenced by their knowledge of how the battle is going.

Barber, McGrew, Stewart and Andrews (1979)
Barber and Kaplan (1979)
Kaplan and Barber (1979b)
Barber and Solick (1980)
Thomas, Kaplan and Barber (1984)
Olmstead, Baranick and Elder (1978c,d)

8. Attempts to correlate individual ARTEP subtasks to battle outcome scores have been singularly unsuccessful. Large matrices of correlation coefficients between some 40 to 45 ARTEP subtasks and as many as eight battle outcome metrics usually show weak, non-significant specific subtask to battle outcome

measure relationships. Care must be taken in attempting to interpret and extrapolate from data that, while occasionally of marginal statistical significance, often has been analyzed only by unsophisticated analysis techniques. These data and findings have a high potential for reflecting a strong halo effect, and quite often show a negative relationship that implies that better subtask performance results in worse battle performance.

Olmstead, Baranick and Elder (1978a, 1978c)
Olmstead, Christensen and Lackey (1973)
Olmstead, Elder and Forsyth (1978)

9. Factor analysis of ARTEP controller subtask scores by effectiveness metric correlations result in either a single factor (battle outcome?) or two rather weak factors (planning the battle and fighting the battle). This result tends to suggest that the subjective nature of ARTEP scores, and affiliated data collection problems - halo effect and the busy controller/evaluator problem - together with limited data, makes battle outcome a poor choice as an effectiveness measure for researchers or command and control performance measurement practitioners. Players and player/controller data when subjected to the same type analysis have developed three factors: 1) gather and analyze information, 2) develop a plan and modify it as required, and 3) implement plan and supervise operations. Suggesting, perhaps, that these persons are less influenced by other factors.

Thomas, Barber and Kaplan (1984)

10. Inspection of ARTEP subtask to battle outcome correlation data suggests that some ARTEP subtasks are generally poorly performed and appear to be somewhat critical to battle outcome. The subtask areas involve fire support, intelligence preparation of the battlefield, operations, logistics and electronic warfare subtasks.

Barber, McGrew, Stewart and Andrews (1979)
Barber and Kaplan (1979)

11. Attempts to increase the measured relationship between staff performance, or organizational competence, and battle outcome scores have been unsuccessful. Early researchers developed a sound method of merging mission accomplishment and resource expenditure data with military judgement to obtain a scale for quantifying battle outcome, but when applied it has not shown acceptably high positive relationships. Research on the battle outcome scores available from simulators (Loss Exchange Ratios, Relative Exchange Ratio, Surviving Maneuver Force Ratio Differential, etc.) show that they correlate highly (e.g., .95) with each other but not with individual staff tasks.

Tiede and Leake (1971)
Thomas (1983)
Thomas and Cocklin (1983)
Thomas, Barber and Kaplan (1984)

12. Organizational effectiveness measures which are command post specific, rather than unit effectiveness measures, have been successfully applied. Decision quality, scoring based on how closely specifics of an OPORD approach a best case or schoolhouse solution, and an indirect measure of OPORD quality (how long does the plan survive and how gracefully does it degrade) have been proposed and successfully used.

Olmstead, Christensen and Lackey (1973)

Krumm, Robins and Ryan (1973)

Robins, Buffardi and Ryan (1974)

Defense Systems, Inc. (1984)

In addition to these quite specific findings past research has resulted in a series of ancillary developments which are of interest and value to command and control performance researchers. Examples of these contributions include: assessment directions for applying the original and later Adaptive Coping Cycle model research approaches, Terminal Performance Objectives and Enabling Objectives for CATTs based - and presumably other simulator based - exercises, several methods for evaluating the quality of an orally presented battalion OPORD, work at both the battalion and brigade level which improves the conduct of ARTEP evaluations and scoring of sub-tasks therein, Terminal and Associated Learning Objectives for brigade command staffs, improved rating forms for Combined Arms Map Maneuver Systems (CAMMS) exercises, and exemplar scales for comparing portions of division OPORDs to presumed best case solutions.

The reader will note in reading the review that the command and control research area is fairly new with the oldest reports dating to the late 1960s and early 1970s. It will also become apparent that conducting good research in command and control is not an easy thing. It is hoped that this review will help researchers and users of command and control research results better appreciate the need for sound practices in conducting and reporting research, and the need to take care in using some of the existing results.

BATTALION COMMAND AND CONTROL RESEARCH

Not unexpectedly there has been considerably more research done on battalion command and control than on higher echelons. The primary reason for the battalion level emphasis is undoubtedly the smaller size of the battalion command post. It is obviously easier to conduct research on a twelve man command organization than on the larger brigade through corps command groups. In addition to the size advantage battalion level research is attractive because the CP - OPFOR - CP cycle is shorter for the battalion CP than for higher echelon command posts. These two fundamental advantages, smaller command group size and the reduced situational complexity resulting from the closeness of the battalion command post to the enemy activity also resulted in battalion level computer based battle staff training simulations being developed before attention was given to higher echelon based simulations. Because of the availability of data from computer aided events in battalion command staff training exercises and the need to verify the training value of these simulations battalion research opportunities were created.

IN THE BEGINNING

In reviewing the research on battalion command and control it is necessary to start from an unlikely beginning. As it happens the research reported by Olmstead, Christensen, and Lackey in 1973 was not only the earliest work reported in the area but it set a standard that has not been exceeded, or matched, in the years since it was reported. As described by the authors:

The study was designed to identify and explore organizational processes critical to effective functioning and to determine how functioning of the processes is influenced by environmental pressures. A framework was developed based upon the concept of Organizational Competence and an Adaptive-Coping Cycle consisting of seven critical processes. Ten 12-man groups of experienced Army officers participated in an eight-hour role simulation of a light infantry battalion engaged in combat operations. Player communications were content-analyzed for quality of process performance, and the organizations' activities were evaluated for military effectiveness. It was concluded that Organizational Competence is a principal determinant of effectiveness, that an organization's ability to respond flexibly to pressures and changes in its environments is related to its competence, and that proficient performance of the identified processes improves effectiveness.

Because the research was conducted in the early 1970s. Olmstead and his associates had certain advantages that have not been available to subsequent researchers. One advantage, of course, was that he had Vietnam experienced infantry officers to draw upon as subjects, controllers, and battalion command and control experts as he prepared for and conducted his research. Another advantage was that in the pre-ARTEP, pre-computer based simulation era in which

they worked the researchers were not constrained by any real or imagined need to embed their research in an ARTEP or computer aided battalion simulation event.

Olmstead and his associates prepared for their research by first developing a model, or conceptual framework, to describe the organization with which they were concerned. They then developed measures for the effectiveness of the organizations products and the internal competence of the organization.

Conceptual Framework

In developing their conceptual framework the researchers drew primarily upon work done by proponents of General Systems Theory. First addressed by von Bertalanffy (1956), the theory views an organization as a system having inputs (resources such as people, material, and information) on which it operates a conversion process (throughput) to produce outputs (products, services, etc.). Later Bennis (1966), based on General Systems Theory, proposed that researchers in the organization arena should be primarily concerned with what he called "organizational health" which he defined in terms of "competence," "mastery," and "problem-solving ability." Bennis suggested that suitable criteria for organizational health include:

1. Adaptability - which coincides with problem-solving ability which, in turn, depends upon flexibility of the organization. Flexibility is the freedom to learn through experience, to change with changing internal and external circumstances.

2. Identity - Adaptability requires that an organization "know who it is, and what it is to do." It needs some clearly defined identity. Bennis says that identity can be examined in two ways: (a) by determining to what extent the organizational goals are understood and accepted by the personnel, and (b) by ascertaining to what extent the organization is perceived veridically by the personnel.

3. Reality Testing - The organization must develop adequate techniques for determining the "real properties" of the environment in which it exists. The "psychological field" of the organization contains two main boundaries, the internal organization and the boundaries with the external environment. Accurate sensing of the field is essential before adaptation can occur.

Clearly Bennis viewed organizations as adaptive systems and felt that the processes through which the adaptation occurs should be the concern of the organizational researchers. Some other researchers of the period certainly agreed. Altman (1966), for example, proposed that organizational researchers could not deal only with the relationships of the inputs and outputs of organizations. He postulated the need for a research strategy, "that peers into the box and attempts to understand the sequential development of performance as it progresses from input to output." Schein (1965), went beyond Altman and sug-

gested an actual sequence of activities or processes used by organizations in adapting to changes in the environment. Schein called this sequence an Adaptive Coping Cycle. The stages in the Adaptive Coping Cycle as defined by Schein were:

1. Sensing a change in the internal or external environment.
2. Importing the relevant information about the change into those parts of the organization that can act upon it.
3. Changing production or conversion processes inside the organization according to the information obtained.
4. Stabilizing internal changes while reducing or managing undesired by-products (undesired changes in related systems that have resulted from the desired changes).
5. Exporting new products or services that are more in line with the originally perceived changes in the environment.
6. Obtaining feedback on the success of the change through further sensing of the state of the external environment and the degree of integration of the internal environment.

Operating from the premises of von Bertalanffy, Bennis, Altman, Schein, and others, Olmstead and his associates elected to define organizational competence using two of Bennis's concepts - adaptability and reality testing - and a third concept, integration, which was derived by the researchers.

Thus, organizational competence was defined in terms of following basic components:

1. Reality Testing. Capacity to test the reality of situations facing the organization - the ability of the organization to search out, accurately perceive, and correctly interpret the properties and characteristics of its environments (both external and internal), particularly those properties that have relevance for the functioning of the organization.
2. Adaptability. The capacity to solve problems arising from changing environmental demands and to act with effective flexibility in response to these changing demands.
3. Integration. The maintenance of structure and function under stress, and a state of relations among sub-units that ensures that coordination is maintained and the various subunits do not work at cross-purposes.

Conceptualization of Organizational Processes

In order to evaluate the competence of an organization it is necessary to define and measure the components that comprise competence. To do this the six stages of the Adaptive Coping Cycle previously discussed were expanded to seven, and a relationship between the three competence components and the seven organizational processes were defined. The seven processes were:

1. Sensing. The process by which the organization acquires information about the external and internal environments.
2. Communicating Information. The process of transmitting information that is sensed to those parts of the organization that can act upon it.
3. Decision Making. The process of making decisions concerning actions to be taken as a result of sensed information.
4. Stabilizing. The process of taking actions to maintain internal stability and integration that might otherwise be disrupted as a consequence of action taken to cope with changes in the organization's environments.
5. Communicating Implementation. The process of transmitting decisions and decision-related orders and instructions to those parts of the organization that must implement them.
6. Coping Actions. The process of executing actions against an environment (external or internal) as a consequence of an organizational decision.
7. Feedback. The process of determining the results of a prior action through further sensing of the external and internal environments.

The relationship of the three competence components to the seven organizational processes were considered to be:

<u>Competence Component</u>	<u>Organizational Process</u>
1. Reality Testing	(1) Sensing (2) Communicating Information (7) Feedback
2. Adaptability	(3) Decision Making (5) Communicating Implementation (6) Coping Actions
3. Integration	(4) Stabilizing

Measurement of Organizational Competence

The manner in which the competence data was treated and scored, though not particularly complex, did involve a considerable number of steps. Basically all the communication data, written, and spoken, recorded during each session was transcribed and sorted so that the material relevant to a particular event or "probe" was gathered into a "probe manuscript." A content analysis was then performed on each communication event in the transcript and each event was coded. The most important coding involved classifying each communication. Individual communications were classified into one of the seven Adaptive Coping Cycle categories and were assigned a value (10, 20, 30, or 40) based on scorer ratings on a four point effectiveness scale (poor, marginal, adequate, excellent). After coding, the data were then sorted to obtain effectiveness scores for process categories, for each probe, etc. Figure 1 shows a schematic portrayal of Olmstead's data collation and analysis process.

Conceptualization of Organizational Effectiveness

In defining their concept of organizational effectiveness Olmstead, and his associates noted that they were dealing with organizations - military organizations - where it was relatively easy to specify the purpose of the organization. Military units are assigned missions and the purpose of the organization is to accomplish these missions. They were able to conclude, therefore, that for the study described in their report, the obvious criterion was whether the organization accomplished its tactical purposes. In defining and operationalizing this concept into a process to measure the effectiveness of a battalion command post, although they did not specifically discuss the matter, they made it clear that they considered that the criteria for command post performance needed to be based on command post specific outputs. They based their measure on the correctness of the decisions the command post made concerning what the battalion should do at the lower level (company, platoon, squad, individual) where the action would finally occur.

Measurement of Organizational Effectiveness

The operationalization of the battalion command post organizational effectiveness measure was a three step process: (1) probe identification, (2) probe response identification, and (3) probe response adequacy evaluation.

Probe identification. The purpose of the Olmstead et al research was to study the functioning of infantry battalion command post operations in rapidly changing combat environments. To provide such a command and control environment, the researchers elected to simulate a Vietnam War light infantry battalion command post involved in a stability operation. In order to obtain realistic scenario events four combat experienced infantry captains were asked to write detailed accounts of events they had experienced first hand, and which they considered to be typical of stability operations. Each account contained descriptions of the locale, physical environment, personnel involved, circumstances leading to the event and the outcome of the event. In preparing the event descriptions each officer was assigned responsibility for a different

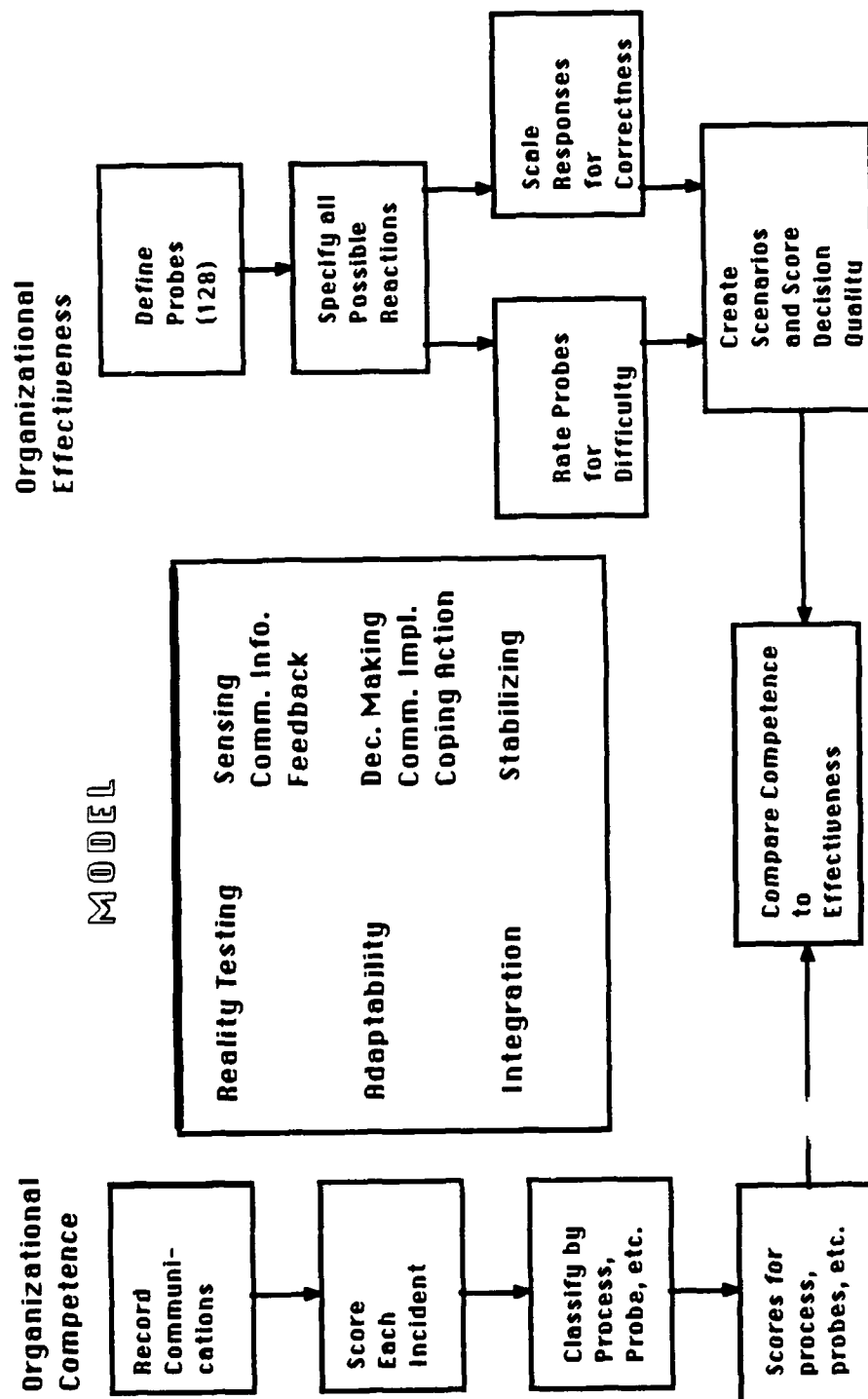


Figure 1. Schematic of the Olmstead et al approach.

subsystem; personnel, intelligence, operations, and logistics. In all 128 events were identified and described in this manner. These events - referred to as probes - became the basis for creating and controlling the scenario.

Probe response identification. After the probes were described a group of seven officers with battalion command post experience considered each probe and listed the possible command post outcomes (including doing nothing) that might result when the probe was presented.

Probe response adequacy. After the possible responses were identified the same seven officers individually, then as a group assigned each possible response a value, 50 through 10, based on a five point scale - highly satisfactory, satisfactory, marginal, unsatisfactory, highly unsatisfactory. It is noteworthy that only 22 times in 1280 probe applications did a response occur that had not been predicted. On these occasions the scorer assigned the value of the response that most resembled the one that had occurred.

In this study the measure of competence was quality - how well processes were performed. The following criteria, reproduced from the original report were used for scoring process performance.

1. Sensing

- a. Accurate detection of all available information.
- b. Correct interpretation of all detected information.
- c. Accurate discrimination between relevant and irrelevant information.
- d. Relevance to mission, task, or problem of all attempts to obtain information about the environment.

2. Communicating Information

- a. Accurate transmission of relevant information.
- b. Sufficient completeness in transmission to achieve full and adequate understanding by recipient.
- c. Timely transmission of information.
- d. Transmission to appropriate recipients.
- e. Correct determination of whether information should be transmitted.

3. Decision Making

- a. Correctness of decision in view of circumstances and available information.
- b. Timeliness of decision in view of available information.
- c. Consideration in the decision process of all contingencies, alternatives, and possibilities.

4. Communicating Implementation

- a. Accurate transmission of instructions.
- b. Sufficient completeness to transmit adequate and full understanding of actions required.
- c. Timely transmission in view of both available information and the action requirements of recipient.
- d. Transmission to appropriate recipients.

5. Actions: Stabilizing, Coping, and Feedback

- a. Correctness of action in view of both the operational circumstances and the decision or order from which the action derives.
- b. Timeliness of the action in view of both the operational circumstances and the decision or order from which the action derives.
- c. Correctness of choice of target for the action.
- d. Adequacy of execution of the action.

The Experiment Itself

Ten 12-man groups of Vietnam-experienced infantry officers, ranging in grade from senior major to first lieutenant, participated in an eight-hour role simulation of a light infantry battalion engaged in combat operations in Vietnam. All inputs into the simulated battalion were made by experimenter/controllers who filled the roles of personnel at brigade, platoon, and adjacent unit levels. Through the use of preplanned and tightly scheduled messages, controllers created a dynamic and realistic situation that provided continual and changing environmental inputs requiring rapid and flexible organizational responses from the simulated units. The simulation scenario consisted of 128 probes (problems) made up of 376 separate input messages. Although activities of the players were uninterrupted, the simulation was designed in four administrative phases, three of which differed in the intensity of environmental pressures, as determined by frequency, complexity, and criticality of probe inputs.

Data collected included: players' ratings of realism, involvement, and pressure experienced during the simulation, and all communications (radio, written, and face-to-face) of members of the simulated organizations. Communications of the players were the source of data for evaluation of both Organizational Competence and Organizational Effectiveness.

The analysis of Organizational Competence included: content analysis of each item of communication according to a system that classified it in terms of 12 descriptive categories and identified the organizational process performed by the item; assignment of a score to each communication item to quantify how well the process represented by it was performed; and the development of group scores for each organizational process, competence component, and competence as a whole.

Organizational Effectiveness was determined by the decisions made when reacting to the 128 probes. Experienced officers examined transcripts of communications concerning each probe and assigned an effectiveness score according to predetermined criteria concerning the potential contribution of the command and control outcome to mission accomplishment. Group Effectiveness scores were summations of scores for the 128 probes.

Results

In the 1973 report Olmstead and his associates discussed their results in ten major subsections. The following are of particular interest in the context of this review: (1) validity of the simulation, (2) competence and effectiveness, (3) organizational processes, (4) process linkages, and (5) the effects of pressure.

Validity of the simulation. As part of the data collection process all players had rated their participation in the command post simulation in terms of how interesting the experience was compared to other battalion CPXs they had participated in, how realistic the problems or probes had been, how likely it was that good performance in the research setting predicted good performance in a real situation, and the degree of player involvement. Results indicated that the players had found the simulated command post exercise more interesting than CPXs, that they considered the problems realistic, and that player involvement had been high. Players also reported that they believed that it was quite likely that the results were predictions of potential real world performance.

Competence and effectiveness. Correlations between the scores of the major variables and subvariables are shown in Table 1. The correlation between overall Competence and Effectiveness (.93) is significant at the $p < .01$ level and demonstrated a strong relationship between the two variables. Indeed since the results indicated that competence accounted for 85.7% of the variance in Effectiveness the researchers concluded that Organizational Competence is the principal determinant of Organizational Effectiveness.

Table 1

Correlations Between Major Variables and Sub-Variables

Variable	Effect	Comp.	Rlty Tstg.	Adapt	Integ.
Effectiveness	--	.93**	.96**	.79**	.11
Competence		--	.94**	.92**	.33
Reality Testing			--	.73*	.10
Adaptability				--	.43
Integration					--

**p < .01; *p < .05. Correlations are based upon eight degrees of freedom.

Table 1 also shows the correlations between the three components of overall Competence and their correlations with Effectiveness. The authors attribute the low relationship of Integration with the other variables to the fact that integration has only one related Organizational Process, Stabilizing, and that the stabilizing process was seldom performed. (Four of ten groups did not perform the process at all.) This situation was considered an artifact of the experimental situation which did not last long enough to require any significant number of stabilizing actions.

The strong relationship between measures was as expected. As was implied in discussing the conceptual framework, the competence components that comprise the Adaptive Coping Cycle are not independent. Rather, a sequence exists in which the quality of each component depends, in part, upon the quality of preceding component. Thus, a relationship would be expected between the seven organizational processes in the Adaptive Coping Cycle with each competence component probably limited in that no matter how well it is performed it cannot correlate higher with Effectiveness or Competence than did the preceding step. Thus, if the organizational processes related to Reality Testing are performed well, the probability of effective performance of the Adaptability and Integration components is enhanced; if Reality Testing is poor, effective performance of Adaptability and Integration will be less probable.

Organizational processes. The relationships between Effectiveness, Competence and the seven organizational processes that form the Adaptive Coping Cycle, are shown in Table 2. Organizational Processes are the fundamental elements of Competence. The seven processes are conceptually different, but not independent, functions that are performed by all organizations. Performance on each process contributed to the ultimate Competence score of each simulated battalion. Accordingly, knowledge of the relationships of each of the processes to Effectiveness, Competence, and to the other processes is important to understanding the dynamics of organizational performance.

For all processes except Stabilizing and Feedback, correlations with Effectiveness were significant beyond the .05 level of confidence. As discussed earlier, the fact that the stabilizing process was not performed by some groups and occurred infrequently in the remaining ones resulted in highly restricted variances which, in turn, produced low correlations with Effectiveness. Feedback, a process that is related to the Reality Testing Component, also was not often performed, with the same result. Obviously in this experiment, Stabilizing and Feedback were not related to Effectiveness. However, because the lack of demonstrated relationship may have resulted from an anomaly in the experimental situation, the researchers concluded that Stabilizing and Feedback are valid processes that are important to Effectiveness in the real world.

Table 2

Correlations Between Effectiveness, Competence, and Process Measures

Variable	1	2	3	4	5	6	7	8	9
1. Effectiveness	--	.93**	.92**	.83**	.70*	.11	.71*	.72*	.03
2. Competence		--	.95**	.72*	.86**	.33	.77**	.77**	.18
3. Sensing			--	.72*	.79**	.32	.58	.65*	.06
4. Communicating Information				--	.30	-.33	.58	.47	-.08
5. Decision Making					--	.63	.59	.67*	.37
6. Stabilizing						--	.14	.17	.49
7. Communicating Implementation							--	.68*	.29
8. Coping Actions								--	.18
9. Feedback									--

**p < .01; *p < .05. Correlations are based on eight degrees of freedom.

Sensing produced the highest correlation with Effectiveness (.92), Communicating Information was second highest (.83), with Decision Making, Communicating Implementation, and Coping Actions somewhat lower and approximately equal (.70, .71, and .72). Thus, those processes concerned with information acquisition and information processing showed the highest relationship to Effectiveness and those concerned with Adaptability were somewhat less strongly related.

To explore these relationships further, a multiple correlation was computed, with the seven processes as independent variables and Effectiveness as the criterion. Neither the obtained R (.97) nor the corrected R (.86) was significant for the limited degrees of freedom (2) that were permissible. However, the obtained Beta weights for the various processes, and the percentage that each process contributed to Effectiveness are of considerable interest. Table 3 summarizes these results.

It is apparent that the five processes (Sensing, Communicating Information, Decision Making, Communicating Implementation, and Coping Actions) contributed to Effectiveness to an important degree. The most striking point was that Communicating Information contributed 43.9% of the variance in effectiveness scores, more than twice the contribution of the next highest process. This finding suggests that Communicating Information made a unique contribution to Effectiveness. The data also show that the other processes, each of which contributed a much smaller amount of unique variance, and indeed, in one case (feedback) made a small negative contribution.

Table 3

Summary of Multiple Correlation Between Processes and Effectiveness*

Process	Beta	Percent Contribution
Sensing	.213	19.3
Communicating Information	.532	43.9
Decision Making	.195	14.0
Stabilizing	.114	1.2
Communicating Implementation	.074	5.0
Coping Actions	.156	11.5
Feedback	-.115	-.4

* The computed multiple correlation (R) is .97; the R corrected for shrinkage is .86. Neither of these relationships were significant.

Process linkages. Olmstead and his associates also reduced and evaluated their data so that they could demonstrate that processes did in fact impact performance of subsequent processes and of Effectiveness. To do this the data were analyzed in the following manner. For each of the five processes that correlated significantly with Effectiveness, group mean values for each probe were computed by summing all pertinent values within the probe and dividing by the number of occurrences. Thus, for every probe, mean values were calculated that represented performance on each of the five processes by each of the ten groups. All mean probe values were then classified as "low" or "high." Values within the range of 10-25 were considered "low" and those within the range of 26-40 were considered as "high." Probe Effectiveness scores were categorized in a similar manner. Classification of scores in this fashion made it possible to evaluate the effects of various high-low combinations of processes upon the performance of other processes and upon Effectiveness.

The following relationships were typical of those presented by the authors to demonstrate the impact of process linkages. When both Sensing and Communicating Information were high, decisions received high evaluations 60% of the time. In contrast, when both Sensing and Communicating Information were low, high-quality decisions occurred only 21% of the time. When Sensing was low and low Communication occurred high quality decisions were made on only 9% of the probes. However, high quality decisions occurred on 60% of the probes where Sensing and Communicating Information were high. This indicates that something more than just good Sensing and Communicating relevant information is required for good decisions, perhaps good judgment or decision-making skills also play a role. The data suggest, however, that high-quality Sensing and Communicating Information skills make effective decisions possible and that, without them, good decisions are much less likely.

A series of evaluations of the linkages led the authors to conclude that their data demonstrated that as the Adaptive Coping Cycle proceeds each step, or process, depends on the quality of the preceding steps.

Pressure. The research design also provided data to determine the impact of increased command post activity requirements on command post performance. Figure 2 shows mean probe Competence scores for the high and low effectiveness groups plotted against activity level. The data show that the relationship between Competence and Effectiveness includes differences in the gradients of Competence degradation between Phases 1 and 2, and differences in the extent of their recovery in Phase 3. Competence deteriorated for both groups during Phase 2. However, for the High Effectiveness groups, the degradation in Competence amounted to an average of 11.3 points per probe, whereas mean scores for Low Effectiveness groups decreased by 22.1 points. Obviously, the increase in pressure that occurred in Phase 2 affected the Competence of the Low groups much more than that of the High groups.

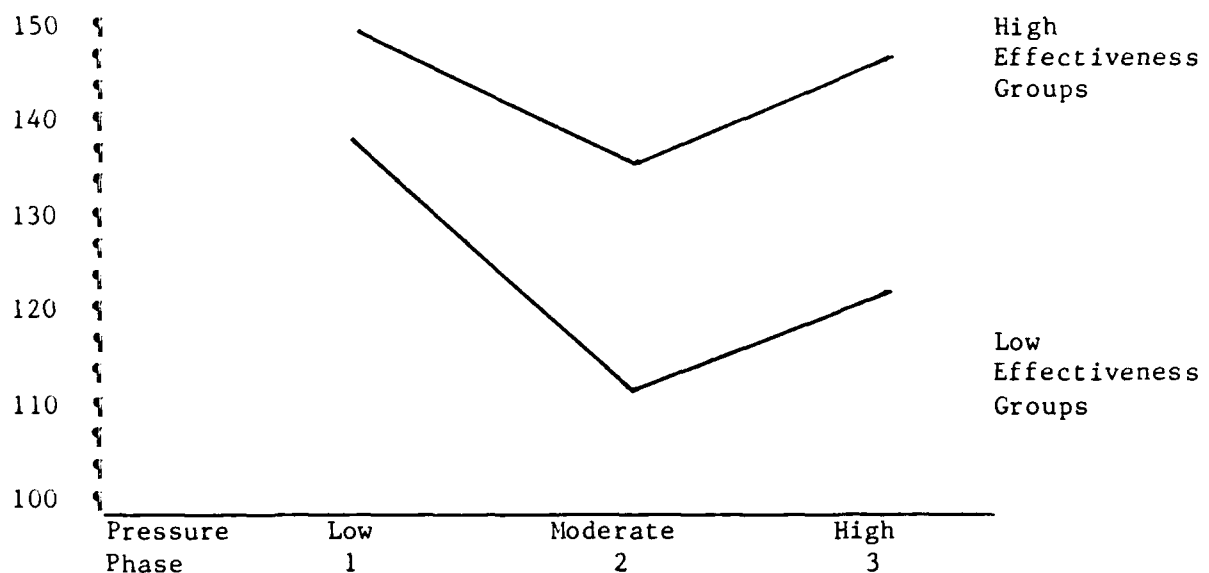


Figure 2. Mean probe competence scores for high and low effectiveness groups under environmental pressure.

High Effectiveness groups recovered Competence in Phase 3 to within about three points of their original Phase 1 level, despite the much more intensive High Pressure condition. Low effectiveness groups also recovered but did not recover quite as much as the more effective groups. Under high pressure, these groups continued to function at a greatly reduced level of Competence and did not approach their original performance. It seems that the capacity of an organization to adapt to rapid or drastic changes or to increased pressure in its environments depends, in large part, upon its ability to adequately perform

the organizational processes that comprise Competence. Hence, the researchers concluded that the quality of process performance is a major determinant of the adaptability of organizations.

Implications of the Olmstead et al. Research

The research described above provides a basis for a series of conclusions relevant to further command and control research, and to command and control performance measurement. The foundations of their research, a model to describe and dissect organizational competence and a well conceived organizational effectiveness measure, enabled the researchers to obtain highly interesting insights into the command and control process.

The research supports the concept that the Competence of an organization can be described in terms of critical processes each of which must be considered an essential contributor to the organization's effectiveness. Five of the seven processes defined by the researchers, Sensing, Communicating Information, Decision Making, Communicating Implementation, and Coping Actions, were significantly related to Organizational Effectiveness and to one another. Although the other two processes, Stabilizing and Feedback, were not significantly related, it is reasonable to accept the authors' view that the nature of the experimental situation was such that these processes did not have an opportunity to occur.

Certainly the research demonstrated that the processes that comprise Competence subsume most of the activities performed by "command and control" personnel in any organization. The results according to the researchers appeared to show the following:

1. Sensing. Results indicated that high-quality Sensing is essential for adequate performance of the remaining processes.
2. Communicating Information. Results indicated that this process makes a unique and significant contribution to organizational effectiveness.
3. Decision-Making. Decisions are affected by the quality of earlier processes and the quality of the decisions then have an impact on later Coping Actions, Stabilizing, formal Sensing Actions, and Feedback.
4. Stabilizing. Results concerning this process were inconclusive. The researchers concluded that further tests of its contribution to the conceptual framework are needed.
5. Communicating Implementation. The activities of individuals who relay instructions between the original decision-maker and the individual who ultimately implements the decision are of particular importance in this process which includes "discussion and interpretation."

6. Coping Actions. Coping Actions are the ultimate determinant of effectiveness. They are affected by all other processes and, in turn, determine the effect of the organization upon the environment.

7. Feedback. Results concerning this process were inconclusive but, because of its heuristic value, the process should be retained in the conceptual framework until further information concerning its validity can be obtained.

The research also demonstrated that Competence is a quality measure of how well an organization performs its critical processes. When the processes are performed well they assist an organization to be effective. When processes are handled poorly, the poor performance will probably negate many positive effects contributed by efficiency in other areas.

In considering their criteria for Organizational Effectiveness the researchers also adopted a sound solution. It was noted earlier that, although the difference between the narrower battalion command post effectiveness and the more global effectiveness of the battalion as a whole was not specifically addressed, the researchers selected an effectiveness measurement process that was command post specific. Thus, they avoided the problems that later researchers faced when they attempted to correlate command and control factors to data where items such as the actions of lower echelon commanders, controllers, simulation rules, replays, and other extraneous factors had an impact.

It must also be noted that although the results of this study are of great value to command and control researchers and to the knowledge base from which command and control performance measures need to evolve, the approach does not solve the measurement problem. The data collection process, transcription of recorded oral communications, and extended analysis process do not lend themselves to routine application beyond the laboratory world.

RESEARCH HIATUS

The Olmstead, Christensen, & Lackey (1973) report, although published in 1973, had been in progress during the 1970-1972 timeframe. After it was completed it was not until 1978 that another battalion specific command and control research effort was reported although in 1977 Olmstead published a brigade study that will be discussed later. There is no certain way to determine why this occurred but it seems, looking back, that during this period there were three developments occurring that could have served to hamper or redirect research in command and control.

The most important would appear to be the movement toward computer supported war games and battle simulations that began in earnest about the time Olmstead and his fellow HumRRO researchers were conducting their research planning and data collection events in the opening years of the 1970 decade. A second thrust in the 1970s that impacted command and control research was the replacement of the Army Training Test (ATT) methodology for measuring individual and unit performance by the Army Training and Evaluation Program (ARTEP)

concepts. A third undercurrent was that in the early-to-mid 1970s there were a series of events that impacted the funding lines and availability of researchers.

Movement to Computer Support

It is probably impossible at this point to determine who first advanced the idea that computer technology should be applied to the exercises, war games and/or other command post activity drivers being used to train and exercise command posts. Shubik in his 1975 book traces the development of an Army model, Carmonette, from its beginnings in the 1950s as a computer aided game called Tin Soldier to a Version IV battle simulation in 1970. (Carmonette stands for Computerized Monte Carlo Simulation). The four Carmonette versions were operational models of combat situations at the battalion and company level. Thus, in the two decades preceding the 1970s the Army milieu included computer aided lower echelon war games and multiple versions of a ground combat model; Carmonette I which integrated tank and antitank operations, Carmonette II which added infantry operations, Carmonette III that added armed helicopter operations, and Carmonette IV which introduced detailed features of communication and night vision. The timing of the development is shown in Figure 3.

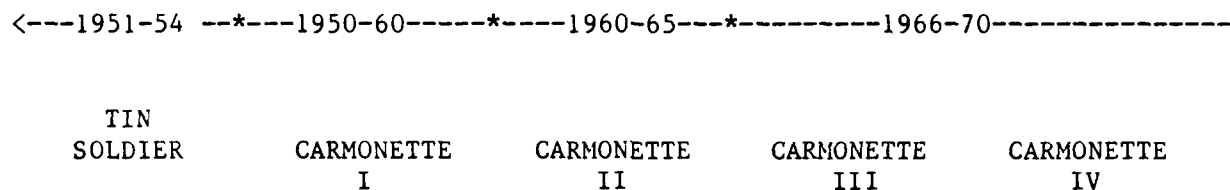


Figure 3. Development of Carmonette.

It is not surprising, therefore, that as the decade of the 1970s began the Army had the expertise, and the will, to move toward the development and application of computer aided games and simulations for use in the command and control training process. Figure 4, by R.E. Solick (Personal communication, October 1987), shows the general geneology of the present family of command and control training systems. By the time Olmstead reported his first research on battalion command and control processes the Computer Assisted Map Maneuver System (CAMMS) was under development at Fort Knox, and the Combined Arms Tactical Training Simulator (CATTS) was being developed at Fort Benning. By 1976 responsibility for development of CATTS had shifted to Fort Leavenworth and the decision had been made to move the technology on to a new, expanded simulation to be known as the Army Training Battle Simulation System (ARTBASS). Meanwhile CAMMS became MACE (not an acronym), which later spun off the Battalion

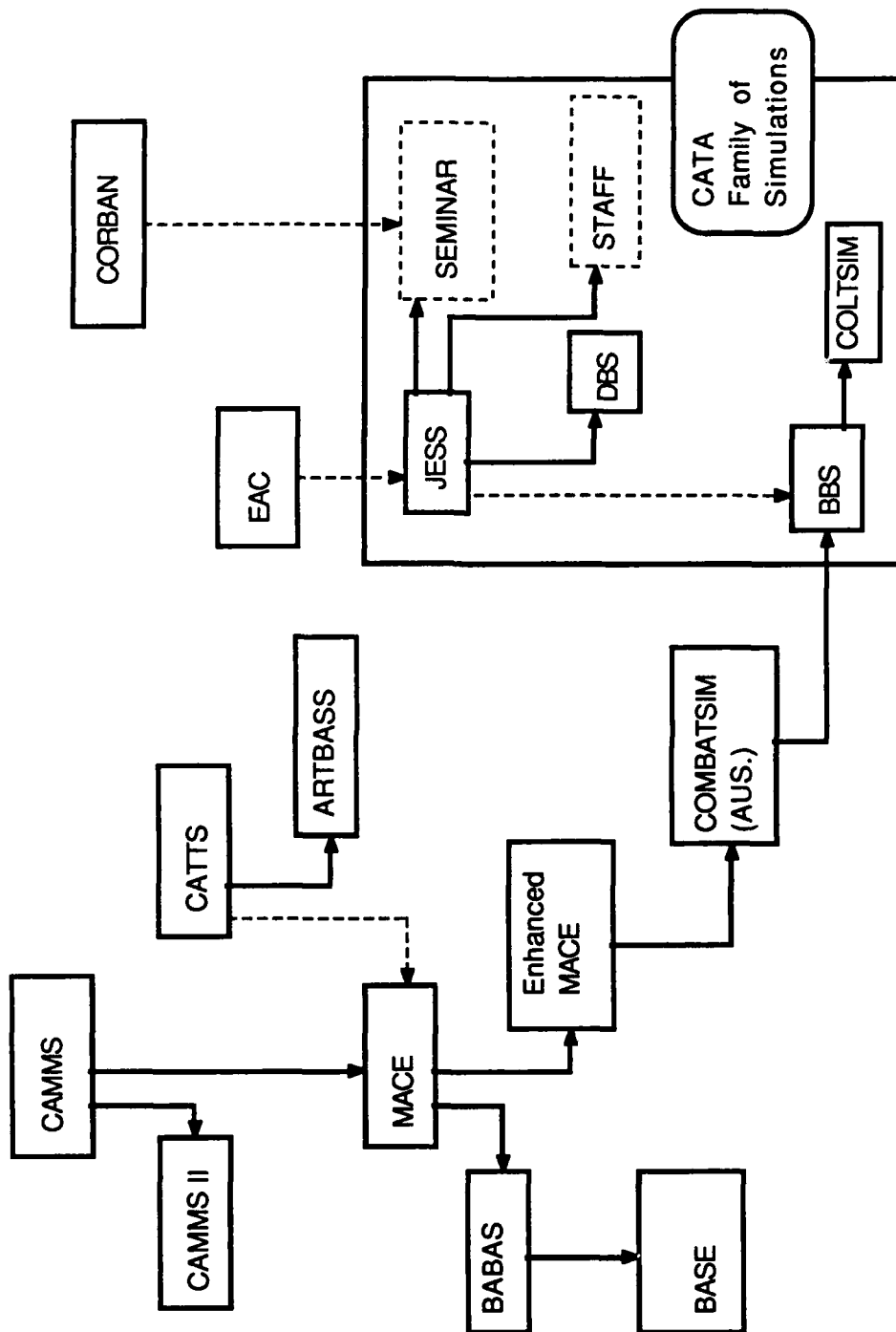


Figure 4. Family tree of Army C2 training systems.

Automated Battle Simulation (BABAS), which by the 1984/5 time frame begat BASE - the Brigade Automated Simulation Exercise. Meanwhile MACE was enhanced and then a version for use by the Australian Army was created.

All this activity suggests quite rationally: that the CAMMS, CATTS, MACE, ARTBASS effort occupied many of the intellectual and money resources of the Army during the early and mid 70s; that research resources were needed to help evaluate the training contribution of the evolving systems; and perhaps even that waiting for development of potentially valuable research tools encouraged prudent researchers to postpone some research.

The ATT to ARTEP Transition Period

The mid-1970 period also saw a transition in the manner by which command and control processes were measured. By 1972 the Army had decided that the Army Training Test (ATT) concept that then existed was unwieldy and inappropriate when applied to group or unit tasks. The bulk of the measures then available for ATT use at these levels had been obtained by extrapolating techniques that had originally been developed to evaluate individual tasks. This had led to evaluation procedures that were excessively detailed and much too subjective when used to evaluate the complex and dynamic skills required for team performance measurements.

By the time Olmstead and his associates had published their 1973 report the movement toward a new, more versatile approach - ARTEPs - had already begun. By 1976 coordinating drafts of command and control staff ARTEP material were being circulated and soon battalion (1977), brigade (1978) and division (1978) ARTEPs had been published for the respective command groups and staffs. Again the situation was such that it may have been prudent to wait the arrival of the new system before funding, or undertaking, new and expensive research. But with the arrival of coordinating draft versions of the new ARTEP system this obstacle was removed, indeed, replaced by a certain urgency to evaluate how well the new system performed.

Research Resources

There were also perturbations in the flow and location of the resources, both financial and human, that were available to do command and control research. Without dwelling too long on the topic it can be noted that in 1972 the Army Research Institute (ARI) was formed by merging two other Army laboratories and given a larger mission that included training research. At about the same time the Human Resources Research Organization (HumRRO) lost, or gave up, its role as a Federal Contract Research Center. The resulting migration of responsibility for some command and control research, and the movement of personnel as some HumRRO functions and people moved elsewhere, including to ARI, worked to create a choke point in the advance of command and control research.

OLMSTEAD REVISITED

The model developed in the earlier section formed the basis for a later series of studies by Olmstead and his collaborators. Shortly after the first report on the development and initial application of their methodology they applied their model to two non-military organizations. In these studies survey techniques were used to evaluate Organizational Competence in 31 social service and rehabilitation agencies nationwide (Olmstead & Christensen, 1973) and in an additional 17 social service agencies in nine states (Olmstead, Christensen, Salter, & Lackey, 1975). In both studies, very strong relationships were found between agency Effectiveness and Organizational Competence, each of the components of Competence, and each organizational process. Effectiveness was measured in terms of agency productivity and judged quality of agency performance.

The results of these two studies led the researchers to conclude that they had confirmed the validity of their conceptual framework as a viable approach for analyzing and understanding the performance of complex organizations. They noted that the results were concrete demonstrations of the importance of the processes subsumed under Organizational Competence as determinants of the effectiveness of organizations, of the relative contributions of the various processes, of the systematic relationships that exist among the processes, and of the ways in which change and pressure influence the performance of organizations. They concluded that it was apparent that Competence played a major role in the performance of organizations and warranted attention in efforts to improve effectiveness.

What appears to be the last research using the original Organizational Competence/Organizational Process model was reported by Olmstead, Elder, and Forsyth in 1980. The purpose of this research effort was to determine if Organizational Effectiveness Staff Officers (OESOs) could be trained to observe and assess the performance of battalion CP staffs and provide feedback of these observations to unit CP personnel. Clearly the method Olmstead, Christensen, and Lackey had used in their 1973 research could not be applied to the measurement of battalion CPX performance unless a quicker way was developed to obtain data on the quality of the performance. The solution adopted was to use a much simpler method in which the observers rated the observed command post for each of the seven organizational processes on a four point scale. The research proceeded in the following manner.

Method

Data were collected in association with Cardinal Point II a training exercise conducted by the 8th Infantry Division during July and August 1978. In Cardinal Point II each battalion was involved in an 11 day exercise consisting of a combination Field Training Exercise, a battle simulation event (PEGASUS), and live firing. The reported data were collected during the battle simulation portion of the exercise.

PEGASUS is a two-sided manual battle simulation that makes possible a training situation in which a battalion command group interacts with controllers playing superior unit levels and with "table controllers" playing friendly unit company commanders. The exercise is planned and directed by an Exercise Director. Activities of controllers are supervised by a Chief Controller. Table controllers are supervised by a Chief Table Controller. Using a control map to depict disposition and movement of forces, friendly table controllers maneuver their units according to instructions from the battalion command group to engage in combat with enemy units maneuvered by Opposing Force (OPFOR) controllers. Engagement outcomes are determined by manual computation using combat results tables provided specifically for use with PEGASUS. Play is activated by an operations order issued by brigade. In the order, a mission is assigned and typical intelligence and other information is provided. Initial friendly and OPFOR unit strengths may be varied according to the training plan and, therefore, differing force ratios may be played. Thus, PEGASUS is a flexible, two-sided, free-play battle simulation which provides dynamic and realistic opportunities for battalion command groups to experience and practice required command and control activities.

Each iteration required four days to complete. Participation of the command groups was continuous, day and night, during each four-day period. Within that period, four modules were completed, with each module consisting of one or more types of combat operations. Within each module, the sequence of activities was as follows:

1. Warning Order from brigade.
2. Fragmentary Order from brigade.
3. Task force orders, terrain reconnaissance.
4. Battle simulation (approximately six hours).
5. Critique.
6. Warning order from brigade for next module.

With minor deviations, the types of operations covered by the respective modules were the same for all units. Similarly, the sequence in which the operations were executed was the same. On the other hand, specific events within a module varied considerably between units because PEGASUS is a free-play simulation and OPFOR players were free to insert special problems, such as chemical attacks, to create particular training situations.

Role of Organizational Effectiveness Staff Officers

Three days before the battle simulation phase of Cardinal Point II, researchers met with the two 8th Division OESOs for the purpose of training them to observe and assess performance of battalion command groups. Two days were devoted to the training. Prior to beginning the training the OESOs studied the 1973 Olmstead, Christensen, and Lackey report, described in a previous section, in order to become familiar with the conceptual framework of the Competence processes developed in that report. The training consisted of review and elaboration of the fundamental concepts; analysis of the PEGASUS configuration for Cardinal Point II and its implications for operational definition of

the processes; revised definitions of the processes for application to single-level organizations, i.e., to command groups alone; procedures for observing and identifying the processes within command groups; procedures for assessing the quality of process performance; and recommendations for providing feedback of observations. In addition, data collection forms and procedures for recording process ratings were reviewed. General questions to be addressed in assessment of process performance are shown in Appendix A. Definitions of the processes and criteria for identifying and assessing them are shown in Appendix B.

An OESO remained with a command group continuously throughout the four-module cycle of the battalion's participation in the battle simulation. Upon completion of each module, the OESO rated performance of the command group on each Competence process according to rating scales developed for that purpose. Then the OESO met with the battalion commander and reported the results of his observations. The implications of the observations for functioning of the command group were then discussed. This procedure of prompt feedback to the commander enabled him to obtain immediate assessment of the quality of process performance in the command group and, if deemed advisable by him, afforded an opportunity to make on-the-spot adjustments in procedures, policies, and behavior of members of the command group. In many instances, at the commander's discretion, OESOs also reported results of observations to command group members and assisted in analyses of ways process performance could be improved.

Thus, an OESO served as "eyes and ears" of a commander with respect to the quality of performance of organizational processes within the command group and provided a mechanism through which on-the-spot feedback could be made available to the commander. In addition, OESO ratings of process performance were the major source of data for this study.

Data collection. Data were collected on 12 battalion command groups; seven mechanized infantry and five armor battalions. Following completion of each module, OESOs rated process performance and OPFOR controllers rated combat effectiveness of the battalions, as described below. The ratings were collected immediately upon completion. Thus, for each battalion there were four ratings (one for each module) of process performance and four of combat effectiveness.

OPFOR controllers varied their inputs depending upon the tactical situation. In addition, different types of operations were judged to vary in terms of difficulty. To obtain some indication of relative difficulty of the four modules, the Chief Controller and the Chief OPFOR Controller rated the difficulty of each module for each unit on a five-point scale. Mean ratings across units were computed to obtain an index of difficulty for each module.

Process performance ratings. After completion of each module, the OESOs assigned to the respective battalions completed the form shown in Appendix C. The form contains seven items, one for each organizational process, upon which an OESO rated the performance of a command group. Raters used a four-point scale, chosen because it was found in the original study of Army battalions

(Olmstead, Christensen, & Lackey, 1973) that assessers of process performance encountered difficulty in discriminating quality of performance when scales of more than 4 points were used. Since one OESO was assigned to each battalion command group, one set of ratings was produced for each command group for each module.

For each battalion, the OESO ratings of performance of a process constituted a Module Process Score, with a possible score range of 1 to 4. Thus, for an exercise (four modules), a unit could receive, for each organizational process, a minimum score of four and a maximum score of 16. Scores for overall Organizational Competence were computed as the sum of the seven process scores. For a module, the minimum possible Competence score was 7 and maximum was 28. Thus, for an exercise, a battalion's Competence score could range from 28 to 112. A composite competence score made up of the total of the seven process scores was also computed.

Combat effectiveness. Upon completion of each module, OPFOR controllers completed the Controller Rating Form shown in Appendix D. The form consisted of five items addressed to a number of aspects of combat outcomes as reflected in the battle simulation. For all items, raters used five-point rating scales with three anchor points having descriptors specific to the content of the items. The five items rated by controllers were: mission accomplishment, area controlled, resources remaining, loss ratio, and overall combat effectiveness. From these ratings Olmstead and his associates developed two additional "composite" scores. One composite was based on a method described by Tiede and Leake (1971) and used by Olmstead, Baranick, and Elder in a brigade command post study reported in 1978. This score was obtained by adding the scores assigned by controllers for Area Controlled and Resources Remaining for each module. The second composite score was obtained by taking the average of the five separate ratings. Thus, in their analysis there were seven combat effectiveness measures; the five original item scores and the two composite scores.

Results

Tables 4 and 5 show the correlations obtained when the researchers transformed the performance rating to standard scores and evaluated the relationship between various scores using the Pearson r correlation coefficient. Of particular note are the correlations of .67 and .63 between the two Combat Effectiveness scores and the total Organizational Competence scores. These correlations were significant at the .01 and .05 levels, respectively. Thus, a significant and strong positive relationship was demonstrated between the quality of battalion command group organizational process performance and combat outcomes as measured in the battle simulation.

Examination of the relationships between combat effectiveness and individual process performance scores, and between Total Competence and the individual elements of combat effectiveness, suggests further conclusions. Combat effectiveness correlated most strongly with the individual processes concerned with Reality Testing: Sensing, Communicating Information, and Feedback. Of these

Table 4

Correlations Between Organizational Process Scores

	Total Competence	Sensing	Information Communication	Decision Making	Stabi- lizing	Communication Implementation	Coping	Feedback
Total Competence	--	.80**	.81**	.79**	.58*	.64**	.66**	.91**
Sensing	.80**	--	.61*	.39	.08	.56*	.25	.85**
Information Communication	.81**	.61*	--	.47	.63**	.65**	.37	.61*
Decision Making	.79**	.39	.47	--	.76**	.21	.90**	.63**
Stabilizing	.58*	.08	.63**	.76**	--	.19	.60*	.27
Communication Implementation	.64**	.56*	.65**	.21	.19	--	.04	.62*
Coping	.66**	.25	.37	.90**	.60*	.04	--	.54*
Feedback	.91**	.85**	.61*	.63*	.27	.63**	.54*	--

* Beyond .05 level

** Beyond .01 level

Table 5

Correlations Between Combat Outcome Scores and Organizational Process Scores

	Total Competence	Sensing	Information Communication	Decision Making	Stabi- lizing	Communication Implementation	Coping	Feedback
Battle Effectiveness 1-5	.67**	.60**	.49*	.49*	.34	.50*	.41	.62*
Battle Effectiveness 2 & 3	.63*	.58*	.32	.55*	.34	.30	.39	.63**
Mission Accomplishment	.56*	.52*	.36	.38	.21	.59*	.27	.53*
Geographical Control	.62**	.59*	.39	.48	.28	.37	.40	.59*
Resources Remaining	.45	.39	.14	.47	.32	.11	.27	.50*
Force Exchange Ratio	.64**	.48	.74**	.36	.39	.56*	.41	.46
Overall Effectiveness	.53*	.47	.43	.34	.22	.39	.37	.47

* Beyond .05 level

** Beyond .01 level

three, relationships were strongest with Sensing and Feedback, the two processes that involve looking for and obtaining information about the battalions combat environment. The quality of processes performed within the organization, such as decision-making, were less strongly correlated with overall outcomes than were the quality of processes concerned with seeing the outside world.

In another analysis the distribution of standardized Organizational Competence scores was split at the median, and the six highest and six lowest battalions were grouped. Table 6 shows mean combat outcome ratings for the separate components and for both Combat Effectiveness scores. All comparisons were significant at the .05 level (one-tailed test) except the rating of Resources Remaining, which narrowly missed significance. The largest differences between high and low Organizational Competence battalions were on the variables of mission accomplishment and force exchange ratio. The authors concluded that the latter was of particular importance in view of the then current U.S. Army doctrine that called for trading time and terrain, within well-defined limits, for the opportunity to inflict disproportionately high losses on the enemy. Clearly this will work only if friendly forces have the skill to achieve a highly favorable Loss Exchange Ratio. Battalions with higher Competence scores in the present research also had better Loss Exchange Ratios.

Implications

It should be noted that in the 1973 Olmstead, et al study the correlation between Organizational Competence and Organizational Effectiveness was .93. In the 1978 study just discussed the correlation was still significant but was only .67 with one effectiveness measure and .63 with the second. The second Olmstead battalion command and control study differed from the first in two significant ways. One was that the data used to define Organizational Competence were obtained in a different manner. The other was that measures of Organizational Effectiveness were different. Both of these changes were undoubtedly a factor in the difference in the results obtained in the two studies.

Organizational competence measures. The reader will recall that in the earlier research (Olmstead, Christensen, and Lackey, 1973) the seven organizational process measures were created from extensive and detailed analysis of all the communications, both written and oral, which occurred during the battalion exercise. This was a slow, carefully controlled process that took many months. In this study (Olmstead, Elder, and Forsyth, 1978c) the same measures were created by having OESOs provide scores for each of the seven process parameters. Since these scores were based on one persons subjective evaluation they were likely to be less accurate than the carefully developed earlier measures. It was also the case that despite their training on the methodology the experienced OESOs provided ratings that differed, (i.e., one rater was consistently higher) so much that a transformation was required to analyze the data.

Table 6

Comparison of Combat Outcome Scores Between More and Less Process Effective Battalions¹

Effectiveness Dimension	More Effective Process		Less Effective Process		t	p ²
	Mean	SD	Mean	SD		
Mission Accomplishment	3.43	.46	2.80	.24	2.96	≤ .01
Geographical Area	3.39	.50	2.83	.38	2.18	≤ .03
Resource Remaining	2.38	.29	2.03	.39	1.76	≤ .054
Force Exchange Ratio	2.91	.19	2.39	.35	3.23	≤ .005
Overall Effectiveness	3.34	.37	2.96	.36	1.81	≤ .05
Battalion Effectiveness 1-5	3.09	.31	2.60	.21	3.21	≤ .004
Battalion Effectiveness 2-3	2.88	.33	2.43	.31	2.67	≤ .02

¹ Sample size is 12 battalions. Scores shown are exercise (four modules) means for six battalions.² One-tailed test.

Organizational effectiveness measures. In the initial study researchers used an effectiveness measure that was based directly on the CP output in response to specific inputs (probes) to the CP. In this study there was no such direct relationship between the cause and the response. The battle outcome data were much less directly measures of the battalion headquarters performance. At the very least battle outcome data includes variance as the result of the OPFOR players doing their best to assure that the concept, attributed to Von Moltke, that "No plan of action can extend with certainty beyond the first contact with the main body of the enemy," remained true. At worst, in addition to other factors introduced by subordinate unit players, there were probably factors introduced by the war game rules themselves.

Ad Hoc groups versus formal CP staffs. There also may have been a third factor at work in the two experiments. In the first experiment the CPs were not actual battalion HQ staffs. Each of the ten groups was created by assigning volunteers to create the required staffs. In the second experiment the staffs were real. It is likely that the newly created staffs were more variable than the existing, already trained staffs which should tend to be more homogeneous. This factor would also work to reduce the correlations.

In any event in the earlier study the correlation between Organizational Competence and Effectiveness was .93, in the latter study the average was .65 (.67 and .63). Thus, given better data, and perhaps more statistically amenable data, Competence accounted for 86% of Effectiveness. With the less exact measures, and perhaps more truncated data, Competence accounted for only 42% of Effectiveness.

Conclusions

Olmstead and his coworkers concluded that the 1977 research validated the concept that organizational processes defined in the 1973 report were indeed related to combat outcomes and that OESOs could be trained to apply the modified measurement system to battalion exercises. The researchers also indicated that considering the subjective, judgmental nature of the observation and scoring system there was merit to the battalion staffs belief that the OESOs assigned to such duty should be experienced combat arms officers. (To support the implementation of this OESO function Olmstead and his associates prepared training material for the U.S. Army Organizational Effectiveness Training Center (see RB 26-5 dated 1 December 1978).)

CATTS BASED RESEARCH

The movement toward using CATTS as a research vehicle for battalion command and control processes began early in the development cycle under the sponsorship and direction of the U.S. Army Infantry School. The initial project, which began in 1972, had two research objectives. The first objective was to conduct an analysis of decision making in command and control situations in order to identify decision making skills required to command and control ongoing combat operations and then to determine relevant training procedures.

The second objective was to assist in the development of CATTS. A report by Tremble, DeLuca, and Lackey (1975a) summarizes the effort. Two other reports describe the research which developed terminal performance objectives for CATTS trainees and a model for combat order production.

Terminal Performance Objectives (TPOs) for CATTS

The terminal performance objectives developed and reported by Tremble, DeLuca, and Lackey (1975b) were such that they had to meet three requirements. First, they needed to define a training milieu that would focus on group skills that would be applicable to dynamic and varied tactical problems of the type that arise in combat situations. The objectives needed to define a method that taught trainees to handle problems, not just to apply fixed solutions to a fixed set of problems. Secondly, the objectives needed to pertain to tasks involved in conducting ongoing combat operations. And, no less importantly, the objectives needed to support CATTS as a team trainer. There was also a fourth requirement implicit in the developmental nature of the CATTS system: the objectives had to be supportable by the hardware and software that would characterize the developmental model of CATTS that was going to be available to test the performance objectives.

The researchers ultimately developed 15 Terminal Performance Objectives. They then further subdivided the TPOs and, described them in terms of 33 Enabling Objects (EOs). Each EO was supported by a definition of the conditions under which the action would be required and standards for evaluating the action. The Performance and Enabling Objectives are listed in Appendix E.

A Model for Combat Order Production

The two HumRRO reports discussed above were accompanied by a report by Lackey, DeLuca, and Tremble (1975) which developed a model of how a combat order was produced by a battalion command staff. To obtain data for model development the researchers first evaluated the decision making literature, read appropriate military documents and discussed the process with experienced military officers. The researchers then content analyzed 22 of the communication transcripts prepared for the previously discussed Olmstead, Christensen, and Lackey (1973) research on organizational competence. These data were used to prepare a order production process model which was then tested by a team of researchers who observed a battalion CP staff planning and executing defensive, offensive, and airmobile tactical missions during a battalion Operational Readiness Training Test (ORTT).

The basic model is shown in Figure 5. The model described order production activities in terms of five components: (a) functions, (b) processes, (c) categories of activities, (d) inputs, and (e) outputs or results of order production performance. Each of these components were reviewed and the processes and categories of performance were presented in detail.

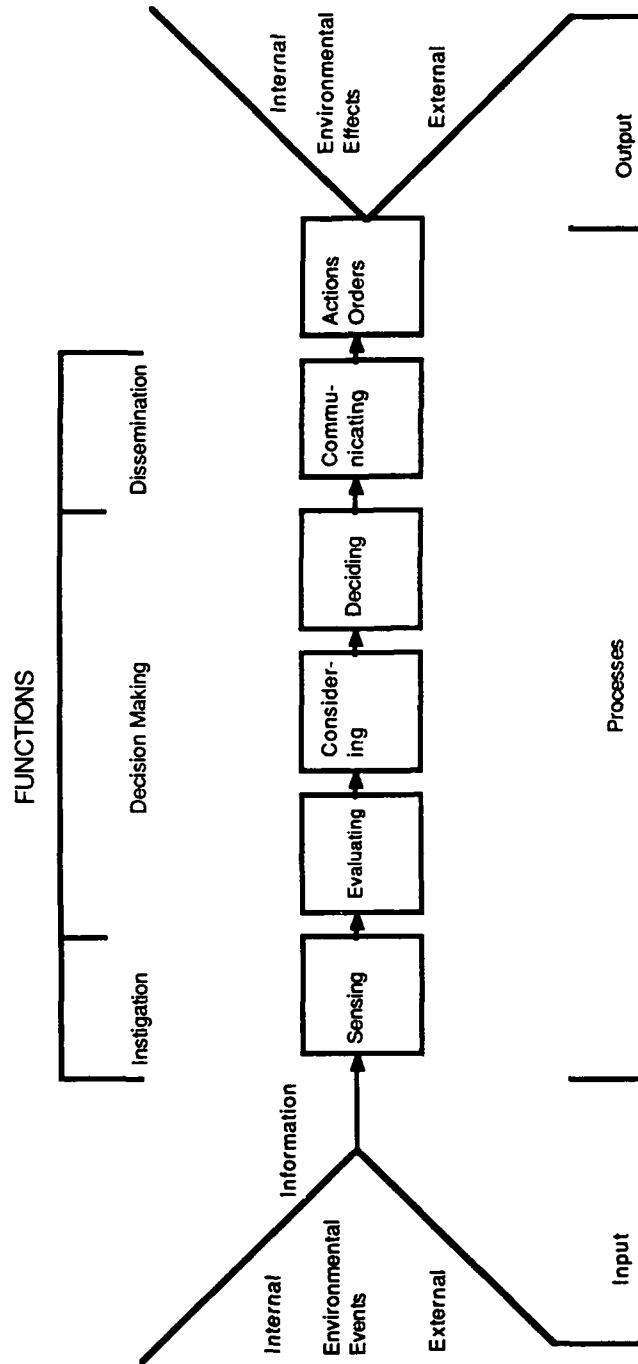


Figure 5. Model of the production of a combat order.

Functions. According to the model, three functions are substantially fulfilled when an order is produced. The first function, referred to as Instigation, involves the recognition or detection of conditions that initiate behaviors leading to decisions. Conditions that will instigate decisions are characterized by a discrepancy either between the current and desired conditions or between projected and desired conditions. A discrepancy can arise from either or both of two sources: it can be created by new orders from higher headquarters or it can result from something that occurs in the battalion's internal or external situation.

The second function, Decision Making, consists of those activities involved in the identification and eventual selection of a course of action for dealing with the discrepancy. Decisions about actions for coping with discrepancies do not necessarily lead to orders. Orders are produced only if a decision is encoded into media appropriate for communication and the encoded decision is then communicated. The third function, referred to as Dissemination, involves those activities that result in the communication of the decision.

Processes. The three functions are accomplished by performance that represents five processes: Sensing, Evaluating, Considering, Deciding, and Communicating. The process of Sensing subsumes those performances involved both in acquisition of information and in the detection of a discrepancy based on that information. The model suggests that in terms of decision making and order production, sensing activities continue until an instigating condition has been recognized.

After the detection of a discrepancy through Sensing, performances regarding the discrepancy center on the function of Decision Making. Three successive processes are involved. The first is Evaluating. The activities represented by this process involve an analysis of the discrepancy to determine its criticality to mission accomplishment. Based on the results of this analysis, activities represented by the process of Considering are undertaken. These activities result in the identification of courses of action that, if enacted, would likely minimize or eliminate the discrepancy. The decision maker then enters the Deciding process and makes the necessary comparisons between alternatives and selects the most appropriate course of action.

The third function, Dissemination, is achieved through activities subsumed by the process of Communicating. These activities include selecting the means for transmitting the decision, coding information, distributing the decision as an order or an action, and reviewing the transmitted order to obtain feedback about its receipt and current appropriateness. The model suggests that the processes are performed in sequence: Sensing, Evaluating, Considering, Deciding, and Communicating and that as the cycle is completed, or aborted, the sensing activities are re-initiated.

The authors also present flow charts that define the activities that occur as each of the five processes occur. These charts indicate which activities represent commander, staff, or command group member actions. The five flow charts are shown in Appendix F.

Operations Order (OPORDER) Based Research

OPORDER planning was considered again in research reported by Archer, Fineberg, and Carter in 1984. In their project data were obtained on five battalion command groups as they underwent training on CATTs. Two data sets were obtained on each battalion: one from the first day CATTs exercise and a second data set from the fourth day exercise. The battalion command groups involved were active Army units. The two exercises consisted of delaying actions sited in the Fulda Gap area of Germany. The exercises differed in location, one in the north, one in the south, and in difficulty. In an effort to minimize these effects three battalions saw the north exercise first and two saw the south exercise first.

Two sets of measures were developed for each of the ten exercise sequences: planning effectiveness measures and battle outcome measures. The planning effectiveness measure was developed by having two experts, a retired Infantry Brigadier General and a retired Field Artillery Colonel view video tapes of the OPORD briefing process and independently evaluate the briefers presentation on three factors: clarity, briefers indication of importance, and the quality of the proposed solution. Rater judgments were made for each of 15 areas. The areas are listed and described in Appendix G which also shows a sample of the Rating Dimensions/Scales form used for each of the 15 information items. For their analyses the researchers apparently used only the clarity and the quality of the proposed military solution. Table 7 shows the planning scores obtained for the ten exercises after the two rater scores were averaged, the resulting 30 scores were weighted based on a criticality factor and, apparently, other manipulations were performed on the data.

Table 7

Overall Planning Scores by Battalion and Day of Exercise

	Day 1	Day 4
Battalion 1	487.2	387.9
Battalion 2	521.3	391.2
Battalion 3	738.5	795.3
Battalion 4	602.6	572.3
Battalion 5	826.4	974.5
Mean	635.2	624.2
Standard Error	64.0	115.0

It is tantalizing to observe that the mean planning scores went down as the units were exposed to CATTS over the four exercise cycles. However, since the difference was not statistically significant the data only justify the conclusion that the quality of planning, as measured by the oral presentation of the plan, did not change from the first to fourth training exposure.

Performance was measured by using battle outcome measures available from the CATTS exercise scores. One measure was based on the Universal Battle Equivalent (UBE) which measures changes in force strength and remaining force strength by evaluating the number of weapons remaining after the battle. A summation of weapons strength is done after each weapon is multiplied by a weighting factor. For example, a truck is weighted 1, an AKM rifle 2, an M-60A1 tank 73, etc. A Relative Loss Exchange Ratio (RER) is then determined from this data by dividing the percent of enemy losses by the percent of friendly losses.

The other two measures were subjective judgments of how well the battalion command staff fought their battle. One judgment was made by the Chief Controller, the second by one of the consultants who assisted in the research. These ratings were made on a nine point scale (poor to outstanding). The raters correlated .92 with each other and .62 and .69 respectively with the RER score. Table 8 shows the RER scores for the ten exercises. The authors indicate that for later analysis the scores were adjusted by adding .437 to the Fulda North scores to compensate for the fact that the Fulda South exercise was concluded to be the easier of the two.

In point of fact the purported .437 difference does not appear to exist in the North Fulda vs South Fulda scores. The difference is only .086. The difference between the day 1 scores and the day 4 scores is, however .431. The adjustment problem - What was adjusted? - casts some doubt on subsequent analyses. It is probably safe to accept the authors statement that "When all ten training exercises were analyzed together there was little significant evidence of a relationship between planning and performance." This should not be interpreted to mean that planning was shown to be of no importance. It may be that the data reflect the fact that the researcher's data collection and subsequent analysis process does not appear to have been adequate enough to identify any meaningful relationships that may exist.

The primary value of this research was in the development of a method for evaluating and scoring the content of the OPORDER as it is briefed at the battalion headquarters prior to the beginning of the simulated battle.

Table 8

Relative Loss Exchange Ratio Scores

Battalion	Day 1	Day 4	North	South
1 - Fulda North	.811		.811	1.480
Fulda South		1.480		
2 - Fulda South	.777		1.403	.777
Fulda North		1.403		
3 - Fulda North	1.148		1.148	1.470
Fulda South		1.470		
4 - Fulda South	.901		1.137	.901
Fulda North		1.137		
5 - Fulda North	1.229		1.229	1.530
Fulda South		1.530		
Mean	.973	1.404	1.146	1.232

Methodology for Assessment of the Planning Process (MAPP)

A CATTIS based study related to the work of Archer, Fineberg, and Carter discussed in the previous section was reported by Metlay et al. in 1985. Metlay and his fellow researchers applied a method of evaluating the performance of decision-making groups that had been developed and applied in research at Hofstra University (Donner, 1981; Frank, 1981; Evangelista, 1982; Metlay, Halatyn & Liebling, 1983; and Otis, 1984). The approach was such that it could be applied by direct observation or by review of videotaped material. The study is closely related to the Archer et al work just discussed because it apparently used the same or similar video tapes. It is also related because although it dealt with the entire planning process it looked specifically at the OPORD presentation.

Metlay and his associates (1985) developed an analytic model of the battalion staff battle planning process based on the military decision making process model that is widely circulated in the military literature. The MAPP model, and "standards", are shown in Table 9. The planning process is evaluated in terms of seven items. Item evaluation is easily accomplished since in most cases it involves merely checking yes or no to indicate if something did or did not occur. Only in four cases does the observer need to make a judgment of whether a high or low level of some activity occurred. The researchers indicated that interrater reliabilities for two judges evaluating the same video taped material during training varied from .90 to .99.

The authors provide tabular data showing the results obtained when the evaluation scheme was applied to video tapes of the planning and OPORD presentation process. The nature of the data - only two samples of five groups - made it unlikely that statistically significant differences could be shown to

Table 9

Model and Related Performance Standards for the Military
Decision-Making Process

Major Categories and Activities	Performance Standards
Information Exchange	
- Staff Activity Level	High
- Battalion Commander's Involvement	High
Mission Analysis	
- Mission Objectives	Yes
- Updates	Yes
- References to Planning Process	Yes
- Obtains Information	Yes
Staff Estimate	
- Occurrence of a Staff Estimate	Yes
- Number of Briefers	Mean = 4.5
- Duration of Staff Estimate	Mean = 18 minutes
Briefers Should Engage in One or More of the Following Activities:	
- Presents Essential Facts	Yes
- Relates Facts to Mission Objectives	Yes
- Makes Predictions and Inferences	Yes
- Presents Conclusions	Yes
- Calls for Questions	Yes
- Delivery and Briefing Style	Yes
- Audience Involvement	High
Commander's Estimate	
- Commander Presents his Concept	Yes
- Commander Presents his Decision	Yes
Preparation of Plans	
- Staff Activity Level	High
- Reference to Time	Yes
- Battalion Commander's Involvement	High/Low as appropriate
Commander's Approval	
- Number of Briefers	Yes
- Duration of OPORD	Yes
Briefers Should Engage in One or More of the Following Activities:	
- Presents Essential Facts	Yes
- Relates Facts to Mission Objectives	Yes
- Makes Predictions and Inferences	Yes
- Presents Conclusions	Yes
- Calls for Questions	Yes
- Delivery and Briefing Style	Yes
- Audience Involvement	High

exist. The tabulated data suggest that groups score higher on the fourth day of exercises which implies that training on CATTS was occurring. It is also apparent that the rating system is easily applicable in real time situations and that there is no reason to doubt that high and/or yes evaluations represent a more adequate planning process. Also, since the duration of briefings are recorded, intergroup comparisons are possible on this parameter.

It should be noted that the researchers also provided a manual, "MAPP Methodology for Assessment of the Planning Process" to aid personnel using the MAPP technique to evaluate the battalion planning process.

ANOTHER BENCHMARK STUDY?

In a later study Carter, Lockhart, and Patton (1983) discuss command and control research done on battalion command post exercises in the CATTS training facility. The fundamental premise of their work was that researchers could separate procedural behaviors from non-procedural behaviors on the basis of data collected in battalion command posts and it would then be possible to analyze non-procedural performance data in order to obtain measures of "individual and multi-individual" behaviors in order to determine the respective contribution of the two types of behavior to command group behavior.

In order to do this the researchers defined non-procedural behaviors as "those for which procedures for performance have not been established or which are not readily amenable to proceduralization." Examples of such non-procedural behaviors included, "assimilation of information, classification of information as to importance, integration of important information in order to perceive a situation, interpretation of the information for impact on military situations, and responsiveness to situations with estimates, plans, and discussions."

Since experimental data had to be collected on a non-interference basis during training exercises the researchers were essentially restricted to those data which were routinely collected and those which could be obtained without changing exercise conditions. The researchers concluded that the battle outcome data, which were routinely collected, were suitable for use as the effectiveness measure since command post performance was the most significant variable influencing battle outcome. Other data from which staff performance could be analyzed was obtained by making videotapes and voice recordings of the command post activities and by reference to extant written records. Data were collected on five battalion command groups from Army units undergoing training at the CATTS facility. Data from the first and fourth training session were used in the analysis.

The CATTS simulation provides six battle outcome measures: Relative Exchange Ratio, Loss Exchange Ratio, Surviving Maneuver Force Ratio Differential, and the same three measures for battle outcome calculations which include artillery fire but only for OPFOR units. Correlations among the six measures for data collected in this research were very high (.95 to .99) so the researchers used

only Relative Exchange Ratio in their analysis. Recorded command staff activity data were then analyzed and coded by a three (or two) person team and correlations were computed between the Relative Exchange Ratio and the 24 scores shown in Table 10.

The failure to obtain significant results is not surprising. Clearly it is quite optimistic to expect to define a rather complex relationship between such detailed sub-processes and a battle outcome measure when there are so many potential variables between each sub-process and the effectiveness measure. Also there is considerable doubt about the process that developed the data. Interrater reliability, between what seems to have been a two person data reduction and coding team, varied from .10 to .84. The average for the 24 measures was a rather low .53. Also, the researchers had postulated that eight variables, marked in Table 10, would have a curvilinear - an inverted U - relationship with Relative Exchange Ratio measures determined after the battle. It was, therefore, improper for the researchers to test their anticipated curvilinear relationships with the Pearson r statistic which the trained reader will recall is a measure of the linear relationship between data sets.

The researchers also collected what they referred to as subjective or qualitative measures. After each command post exercise the Chief Controller and a consultant assigned to observe the exercise recorded their perceptions concerning how well the battalion staff exhibited the behaviors implied by the seven qualitative parameters shown in Table 11. The controller/consultant scores were apparently averaged and the resulting scores used to calculate correlations with the battle outcome score. These correlations are also shown in Table 11.

The sometimes significant correlations shown in Table 11 may well be deceptive. Since the ratings were made after the exercise, and, hence, battle results were pretty well known, they may include a strong halo effect. If the rather robust correlations do not reflect some data development problem they certainly must suggest that the researchers were able to quantify and judge broad, vaguely defined parameters better than the narrower, better defined parameters listed in Table 10. If this is the case, the planners of the research certainly overlooked an opportunity to collect better, more relevant data. This study adds little to the material previously reported in the battalion command and control area, but it does demonstrate that the difficulties in conducting research in this area are significant and that even ill advised efforts can give what appear on the surface to be good results.

Table 10

Statistical Correlations Between RER and Battalion Data

Variable	Reliability	Correlation with RER	Significance
<u>Data Transfer</u>			
Data Transfer Frequency*	.70	.10	.38
Data Transfer Duration*	.60	-.10	.39
<u>Communication Mode</u>			
Telephone Frequency*	.54	.21	.28
Telephone Duration Ratio*	.86	-.05	.45
Radio Frequency*	.68	.03	.46
Radio Duration Ratio*	.76	-.16	.33
Face-to-face Frequency*	.68	.13	.36
Face-to-face Duration Ratio*	.43	.05	.44
<u>Staff Behavior</u>			
IMI Frequency**	.42	-.00	.49
IMI Duration Ratio**	.41	-.08	.41
Team Frequency	.58	.22	.26
Team Ratio	.73	-.15	.34
<u>Staff Processes</u>			
Seeking Info Frequency	.68	.04	.46
Seeking Info Duration Ratio	.58	-.09	.41
Receiving Info Frequency	.52	-.19	.30
Receiving Info Duration Ratio	.72	-.52	.06
Coordinating Frequency	.19	.08	.41
Coordinating Duration Ratio	.31	.00	.50
Transmitting Info Frequency	.48	.07	.42
Transmitting Info Duration Ratio	.39	-.25	.23
Assessing Situation Frequency	.12	.16	.32
Assessing Situation Duration Ratio	.10	.33	.17
Total Processing Frequency	.64	.10	.38
Total Processing Duration Ratio	.56	-.06	.43

* Curvilinear relationship assumption.

** IMI = Individual and multi-individual.

Table 11

Correlations of Qualitative Measures

Qualitative Measure	Battle Outcome	P Value
Competence	.8732	.027
Reliability	.7461	.074
Functional Integrity	.8091	.049
Flexibility	.6289	.128
Professional Quality	.8425	.036
Personal Demeanor	.7363	.078
Team Awareness	.6139	.135
Overall Battalion Command Group	.7938	.054

WHO KNEW WHAT, WHEN?

Kaplan, in 1980, reported what appears to be the first research specifically directed toward measuring information flow in command groups. Past research, such as the Olmstead et al 1973 work, had often measured the flow of information through the command and control cycle but the data obtained were used as the basis for other measures. Thus, Olmstead measured information flow but analyzed his data to evaluate his sub-processes "communicating information" and "communication implementation". Kaplan's research specifically addressed the issue of how information flowed through a battalion command post. He collected data in such a manner that he could evaluate how information was lost through the process of communicating and remembering command and control relevant material.

Kaplan designed his research to obtain data in three areas. He wanted to determine:

1. How much information is lost during the planning stage?
2. Where is the information lost?

3. How much does individual differences effect the information loss process?

Kaplan conducted his research on 13 battalion command groups who were participating in simulated combat exercises on CATTS at Fort Leavenworth. Each group consisted of seven principal members - the Battalion Commander (BC), S1, S2, S3, S4, Fire Support Officer (FSO), and Air Liaison Officer (ALO) - and generally ten other officers, NCOs and enlisted persons. Two company commanders (CC) also came with the command group to serve in a dual player/controller role during the battalion exercise. The procedure used to measure information flow is shown in Figure 6. Table 12 shows the items that each individual required - or at least should have received - and the person from whom he should have gotten the information. There were a total of 35 items, of which, five were known after the controller prebriefings to only the BC and S3, 15 were known only to the S2, the FSO knew 3 of the items and the S1, S4, and ALO each knew 4 of the remaining 12.

The information items (such as "number of air sorties allocated to division by day", "The initial Fire Support Control Line (FSCL)" were created and assigned such that they were required information for battalion planning activities of the persons to whom they were given by the controllers. It is particularly noteworthy that company commanders (CCs) depend entirely on the battalion staff for their information. The test, administered after the planning - briefing - battle preparation cycle was completed, consisted of 35 questions. The questions were multiple choice and were composed to a format that had one correct answer, two incorrect distractor items and an "unknown" response. Test instructions were that testees should use the "unknown" response rather than to guess.

Figure 7 summarizes the communication and remembering patterns of the command group as a whole. There were 13 battalions, seven direct recipients of the information and 35 information items. The figure shows how information required to do staff or company command jobs flowed and how some non-required information flowed. Table 13 shows the information known after the cycle by the command group itself and by the company commanders. The reader should remember that the "received information" variable and the "remembered information" variable are confounded. As a result although the reader may presume that any staff member would be sure to obtain information that doctrine and the proper planning process defines as required this might not happen. Hence, the percent of recalled item scores most likely represents both forgetting and not having gotten required information.

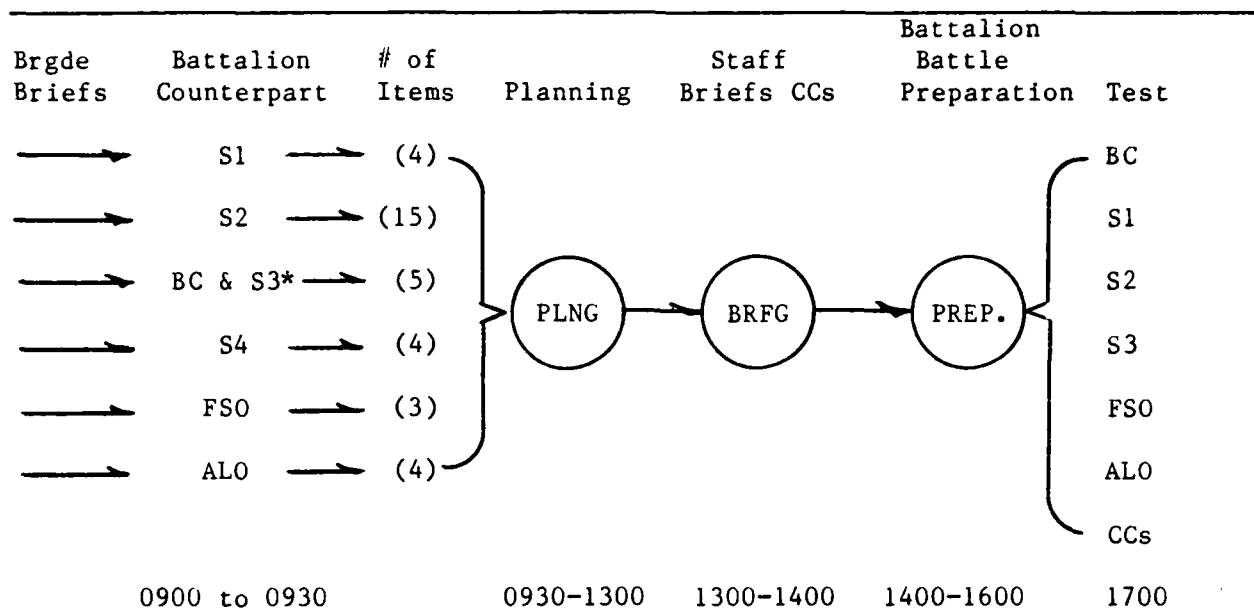


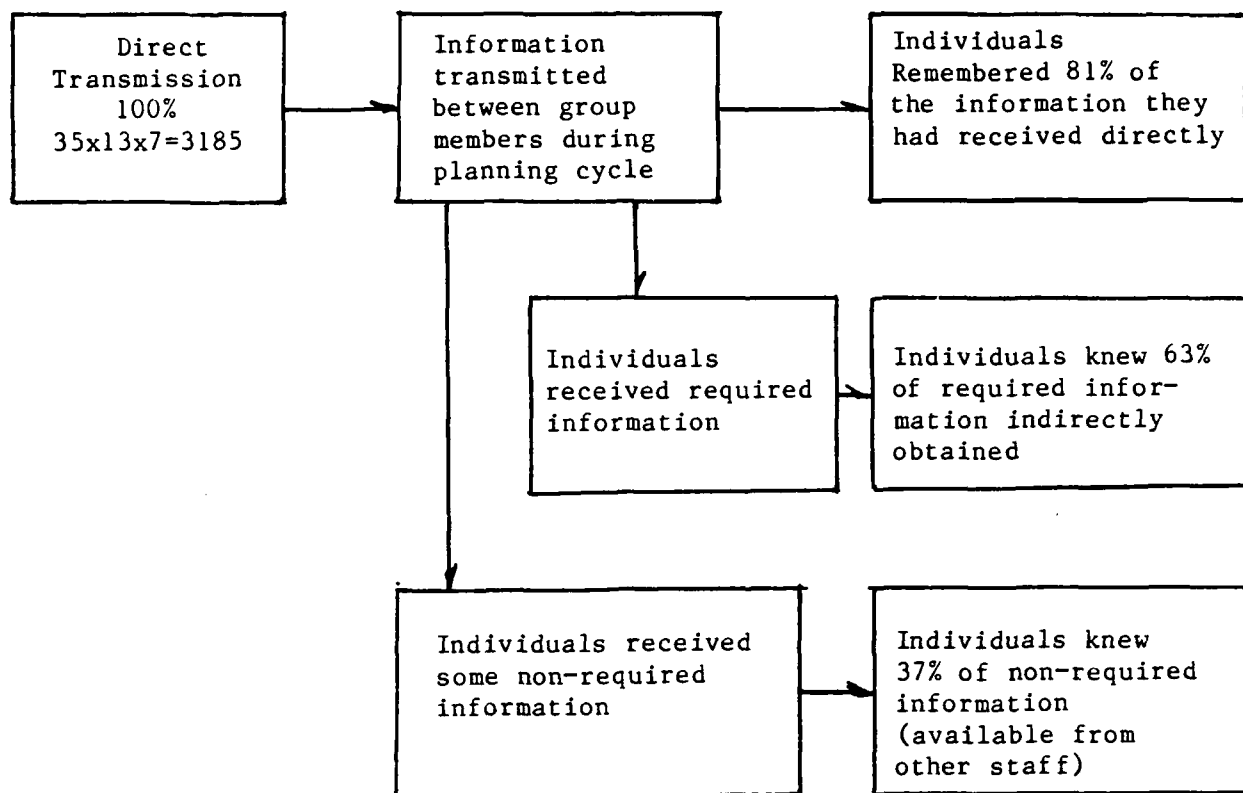
Figure 6. Flow of information from 0900 to 1700 hours.

Table 12

Number of Required Items Each Person Should Have Received and Persons to Whom
Receivers Should Have Sent Information Items

Sender	Receiver							
	BC	S1	S2	S3	S4	FSO	A10	CC
BC & S3*	-	0	3	-	0	3	3	5
S1	3	-	1	2	2	1	1	2
S2	11	3	-	9	3	4	6	8
S4	2	2	2	1	-	0	0	3
FSO	2	0	1	2	0	-	1	1
ALO	4	0	1	4	0	2	-	0
Staff Given								
Total	22	5	8	18	5	10	11	19
Controller								
Given	5*	4	15	5*	4	3	4	0
Total Items								
To Be Known	27	9	23	23	9	13	15	19

* The same five items were presented to the BC and the S3, who were briefed together.



NOTE 1: The fact that 63% and 37% total 100% is only coincidental.

Figure 7. Flow and remembering of information during a CATTS cycle.

Table 13

Intragroup Communication Percentage of Available Required Items Recalled by Each Receiver Averaged Over 13 Battalion Command Groups

Transmitter	BC	S1	S2	S3	S4	FSO	ALO	CC
BC & S3	-	-	69	-	-	79	75	78
S1	42	-	44	37	59	56	63	37
S2	78	58	-	71	72	80	57	66
S4	58	17	50	62	-	-	-	35
FSO	58	-	25	55	-	-	50	41
ALO	65	-	36	60	-	62	-	-

This research provided several provocative results. For example, Kaplan reported that there was a correlation of $-.68$ between measures he developed from his data that represent a transmit-known-information score and receive information scores for individuals. He noted that battalion commanders who transmitted the most information to their staffs received the least information from their staffs. Kaplan postulates that this results from management style and notes that some battalion commanders directed and dominated their staffs to the extent that they interrupted and corrected subordinates as they were speaking - a situation that surely chills the upward communication process. Kaplan contrasts this to commander behaviors that represented a more consultative style involving much more listening and little specific direction. He also noted that a substantial amount of information was lost in the process of communication, (battalion command group members recalled 81% of the information presented to them by their brigade counterparts, and only 63% of the required information available to them from the other members of their own group), that much of the information loss was concentrated in specific intragroup communication channels, some being as low as 17% successful; and that the most important members of the group were listened to the most.

Two later studies, conducted in a similar fashion, have also been reported. Thomas, Barber and Kaplan (1984) reported on a study of five battalion command groups. As part of the study researchers administered a 33 item questionnaire to track the flow of items presented to the BC, S1, S2, S3, S4, FSO and CCs. Data was collected after the first and after the fourth exercise. The results, shown in Table 14, were similar to those just discussed except that day two to day four comparisons were possible. The authors report that improvement was

not statistically significant but state that they believe the improvements would have been greater if battalion staffs had received feedback about their communications behavior.

Table 14

Mean Information Flow Scores for Exercise Days 2 and 4

	Percent	
	Day 2	Day 4
Direct Reception (Brigade to Battalion)	85%	89%
Intra Group Communication	50%	55%
Communications to Company Commanders	47%	56%

Later Thomas, Kaplan and Barber (1984) reported a second study that included a communication evaluation segment. This study, which was also reported in a preliminary version by Thomas, Barber and Kaplan in 1983, involved five battalion command groups undergoing training on CATTs. The questionnaires were administered after the first exercise and after the fourth exercise. Data were collected from six members of the staff (S1, S2, S3, S4, FSO and ALO) who had been briefed by their brigade counterparts and given specific information items. The study differed from the previously discussed research in several ways. One, was that the battalion commander was not questioned. Another difference was that after the first exercise the battalion commander was shown the results of the communications questionnaire and presumably was able to take some action to attempt to improve information transfer. A third difference was that exercise controllers were asked to observe the OPORDER briefing and record information concerning its intent and adequacy.

Results of this study were also quite interesting. An average of 13.2 items were briefed to company commanders in the pretest versus 17.6 items in the post-test. The increase was not statistically significant. However, the average rating of the briefings increased from 4.5 (less than good) in the pretest to 5.9 (half way between good and very good) on the post-test. This improvement was statistically significant at the .01 level ($t = 3.96$). Favorable rater comments on the briefings also increased from 27% in the day one OPORDER presentation to 76% in the fourth day presentation. Table 15, which shows the same type of data as is shown in Table 14, shows the information flow scores obtained from the communication-recall questionnaires given on the first and fourth exercise day. The brigade to battalion change was the same as

in the earlier study (85 to 89). The gain was the same as in the previous study but in this case the difference was statistically significant. This probably results from a reduced variance that in turn was mere happenstance since there were no methodological changes. The other differences, improvements all, were also statistically significant.

Table 15

Effect of Training with Feedback on the Mean Percent of Required Information Received/Remembered

Stage of Communication	Exercise 1	Exercise 4	t
Brigade to Battalion	85%	89%	2.21*
Within the Bn Cmd Grp	51%	66%	4.97**
Battalion to Co Cdrs	42%	67%	4.96**

* $p < .05$

** $p < .01$

df = 4 one-tailed test for matched pairs

SIMULATORS AND ARTEPS

By the late 1970's two of the previously mentioned efforts, the development of ARTEPs and computer assisted battle simulators had progressed to the point where meaningful research could be conducted in situations where both were used. The first battalion level research in this area (Kaplan and Barber, 1979a) compared CAMMS based battalion command and control exercises to CPXs in terms of cost effectiveness. Later CAMMS and CATTIS were both used in conjunction with ARTEP measures in four diverse research projects.

CAMMS versus a CPX

The Kaplan and Barber (1979a) research was conducted primarily to assess the comparative effectiveness and cost of two alternative methods, CAMMS and a conventional CPX, for training battalion and brigade command and control groups in combined arms operations. Both CAMMS and conventional CPXs were intended to achieve the same trainee performance objectives, but developers of CAMMS and

other battle simulations intended to produce systems that would provide a more realistic exercise than a CPX, that would motivate the players more intensely, and that would exercise trainees more thoroughly in the command and control tasks specified in the battalion and brigade ARTEPs. The Army, of course, was also interested in obtaining these advantages but without, at the same time, increasing the cost of the training process.

Method. Kaplan and Barber obtained their data from 50 battalion command groups and 21 brigade command groups after they participated in CAMMS exercises. Data were collected using questionnaires. There were two sets of questionnaires each consisting of two parts. The first questionnaire obtained data concerning how well the CAMMS experience compared to the experiences players had had in CPXs. The second portion of the questionnaire asked for more specific information concerning how well CAMMS exercised specific tasks, taken from the appropriate command group ARTEP, compared to CPXs. Tailored sections were also prepared for the commanders, S1/S4, S2, S3 and FSO. A second set of questions obtained data from exercise directors and commanders to compare the cost of preparing and conducting CAMMS exercises and CPXs.

Results. Kaplan and Barber demonstrated that the cost of preparing a CAMMS supported exercise that provided an opportunity to observe the appropriate ARTEP tasks were 25 to 30% less than those incurred in preparing and conducting an equivalent CPX. The cost differences were both statistically significant and appreciable. The researchers were also able to conclude that CAMMS was considered significantly and consistently more realistic and interesting than a CPX.

The functional areas in which CAMMS enjoyed the greatest advantage were related to preparing and organizing the battlefield; controlling and coordinating combat operations and concentrating combat power (as rated by the S3); and the exercise of command and control (rated by the commander), especially the exposure to the capabilities of enemy weapons systems, facing a thinking enemy, and making decisions under real-time constraints. The principal weaknesses of CAMMS exercises were that they did not produce much stress; they did not exercise the players in security procedures, such as electromagnetic and communications security; and they did not require players to react to special situations, such as enemy jamming, and chemical, biological, or nuclear warfare. The CPX also received low ratings in these areas so the CAMMS weakness were not unique.

A Farewell to CAMMS

The CAMMS training simulation did not move directly into the Army training arena. Indeed it remained up and active only long enough to support three studies in addition to the Kaplan and Barber research discussed in the above section. CAMMS did not actually die, it moved on as part of the CAMMS-MACE-BABAS-BASE development sequence discussed earlier. CAMMS also served, while undergoing its evaluation, as the test bed for a brigade research project, (Olmstead, Baranick, and Elder, 1978c) which will be discussed later, and two

battalion level projects all of which used performance measures developed from ARTEP evaluations.

CAMMS: training effectiveness and research suitability. In the first CAMMS based battalion study that specifically addressed ARTEP tasks as performance measures the researchers (Barber, McGrew, Stewart, & Andrews (1979)) used the five battalion research model already discussed several times in earlier sections. Data were collected on the first and third day of a three day exercise. Two types of data were collected, subjective data based on observer ratings according to the existent ARTEP and objective performance measures. The existing ARTEP (71-2, dated June 1977) described 61 subtasks relevant to the command and control process. Of the 61 subtasks 14 could not be measured because either they could not be observed or they could not be elicited in the CAMMS situation. The remaining 47 were measured subjectively, objectively, or both. In the end 28 subtasks were measured only by the ARTEP process and 19 were measured by the ARTEP checklist and also by a more objective, or at least a more detailed process.

The authors also collected a considerable number of non-ARTEP related measures such as thrust depth, number of decisive engagements, task force losses, OPFOR losses, mission accomplishment, locus of control, intelligence processing, shift and concentration of forces, distance between battle positions, number of battle positions, distance of specified battle positions from the engagement, engagement range of all battles fought and the usual assortment of battle outcome measures. The significance of these measures will be addressed again in a later section.

The researchers attempted to analyze their data by a series of analysis of variance calculations and by extensive correlation analysis. When they correlated their eight battle outcome measures to their 47 ARTEP subtask scores they found that 48 of the 376 values in the resulting matrix were significant at beyond the $P < .05$ level. Table 16 shows the 39 ARTEP subtasks which correlated with five of the eight battlefield outcome measures. Three other measures, Task Force Losses, Decisive Engagements and Area, accounted for only nine correlations and were not included.

In their discussion of these results the authors conclude that Column 1 data can be taken to mean that the smaller the error in identifying the location of the enemy thrust, the better the performance on the respective ARTEP subtasks.¹ This seems to make sense given that three of the four ARTEP tasks concern the planning and coordination of fire support and the fourth concerns integrating CSS into the scheme of maneuver. Plausible explanations for the positive correlation of "communicate/coordinate plans and orders" with amount of error in locating the enemy thrust were described as elusive.

¹Better stated perhaps as "the better the performance on the respective ARTEP tasks the smaller will be the error in identifying the location of the enemy thrust."

Table 16

Statistically Significant^a Correlations with ARTEP Ratings^b for the Five Battlefield Outcome Measures Having the Greatest Number of Significant Correlations

Subjective Outcome						Objective Outcome			
Enemy Thrust		Enemy Losses		Mission Accomplishment		Friendly Personnel Losses		Friendly Equipment Losses	
		1A	.88	1J	.87	1C	.85	1C	.85
1J	-.91	3A	.94	1K	.86	1J	.88	1J	.81
1K	-.85	3C	.94	1L	.87	1L	.88	1L	.81
1L	-.91	3H	.83	3A	.90	3A	.91	3A	.90
3G	.90	6A	.93	3B	.80	3C	.91	3C	.90
9D	-.85	8C	.93	3C	.90	5D	.81	9B	.81
		9B	.90	3G	.90	9D	.86	9D	.80
				3H	.90				
				3I	.80				
				5D	.93				
				6A	.97*				
				8C	.97*				
				9D	.92				

^a All entries are significant at $P < .05$ level.

^b ARTEP tasks and subtasks are shown in Appendix H, the digital-letter combinations refer to that list.

* Significant at $P < .01$ level.

The relation of enemy losses to the seven ARTEP subtasks shown in the second column of Table 16 all seemed reasonable to the researchers. The better one analyzes the mission, determines critical place, organizes for combat, reinforces terrain, modifies scheme of maneuver, concentrates and shifts combat power, and fixes his systems, the greater the losses he is likely to inflict upon the enemy.

Mission accomplishment (Column 3) was the outcome measure having the greatest number of significantly related ARTEP subtasks. That three of these subtasks related to fire support planning was taken as evidence to again confirm the importance of fire support planning. The fact that six subtasks concerned with preparing and organizing the battlefield were also highly related to mission accomplishment, was not seen as surprising. Similarly tasks from the execution portion of the ARTEP, disseminating critical combat information and intelligence along with modifying scheme of maneuver, concentrating/shifting combat power and integrating CSS into scheme of maneuver were highly related to mission accomplishment as might be expected. Columns 4 & 5 provided equally acceptable explanations of the significant relationships.

In general the researchers felt that all the relationships shown in Table 16 were reasonable ones but noted that their conclusions needed to be tempered to reflect the problems associated with attempting mighty analyses with only minor amounts of data. The researchers concluded that most ARTEP tasks and subtasks can be measured on CAMMS and that the system has the potential to be both a training and training research vehicle. Beyond this they concluded merely that the various relationships between their multitudes of data are complex and not conclusive.

Barber and Solick (1980) have also reported on CAMMS based research that incorporated an ARTEP based thrust. There were three goals for their research: to see if CAMMS team training was adequate, to collect data on the implementation and utility of ARTEP measures, and to compare CAMMS evaluators to a special team of evaluators as feedback providers on training deficiencies. Data were collected in Europe on two battalions undergoing a training sequence which consisted of two exposures on successive days to a CAMMS based battle simulation followed, on the two successive days, by an integrated training situation where CAMMS and MILES were combined to produce the training environment. After considerable statistical manipulation of limited data the researchers concluded that their research had demonstrated the feasibility of integrating battle simulation and engagement simulation techniques, that the CAMMS training course was adequate but could easily be improved, and that exercise controllers felt that they had little opportunity to observe performance because their controller duties interfered with observer duties. Full time observers had much more confidence in their ability, and opportunity, to make sound observations and judgements.

CATTS as ARTEP Driver

During the late 1970s the CATTS training battle simulator was also used to evaluate the new ARTEP system and to evaluate the potential of conducting

ARTEPs in battle simulators rather than in the traditional CPX or FTX environment. In the earlier of two ARI research projects Barber and Kaplan (1979a) conducted a study to determine how well the battalion command group tasks defined in Chapter 10 of ARTEP 71-2 (1977) could be exercised and measured in battle simulations conducted on the CATTS simulator. Their project, as defined by the authors, had four objectives:

1. To adapt the Battalion Command Group ARTEP to a computer-driven simulation (CATTS).
2. To identify the ARTEP subtasks on which the performance of incumbent battalion command groups was comparatively weak.
3. To describe the specific behaviors that contribute to inadequate performance of subtasks.
4. To measure the relative criticality of each subtask by determining its relation to overall measures of command group effectiveness.

The ARTEP tasks and subtasks: The battalion command group activities described in ARTEP 71-2 consists of twelve tasks which in turn are composed of 65 subtasks. The tasks and related subtasks are shown in Appendix H. Figure 8 shows a flow model of the twelve tasks.

Procedure. Data were collected from 27 battalion command groups that participated in a simulated defense or covering force operation, and in an attack. The performance of the command groups on the ARTEP subtasks was evaluated by eight observers, seven of whom were also controllers in the exercise. Each evaluator observed specific subtasks, rated the command group's performance on those subtasks, and recorded observed deficiencies. The evaluators also rated the overall effectiveness of individual group members, and of the command group as a whole. Individual tasks were rated on a three point scale: (1) major deviation from ARTEP standard - unsatisfactory, (2) minor deviation from ARTEP standard, and (3) satisfied ARTEP standard. The overall performance rating was made on a five point scale: one of the worst, worse than average, average, better than average, and one of the best. The same five point scale was used by the S1/S4, S2/S3 and five support controllers to evaluate their player positions. Mission accomplishment was also judged by the non-controller monitor who scored mission accomplishment as: no, marginal, or yes.

Results. The researchers reported their results in two areas: relative performance of ARTEP subtasks and the relationship of the subtasks to overall performance. By allocating specific subtasks to the individual controllers who were in a position to observe them, it was possible to evaluate the command group's performance on most of the subtasks in the ARTEP module, and, thereby, identify those subtasks that were highly correlated with measures of overall

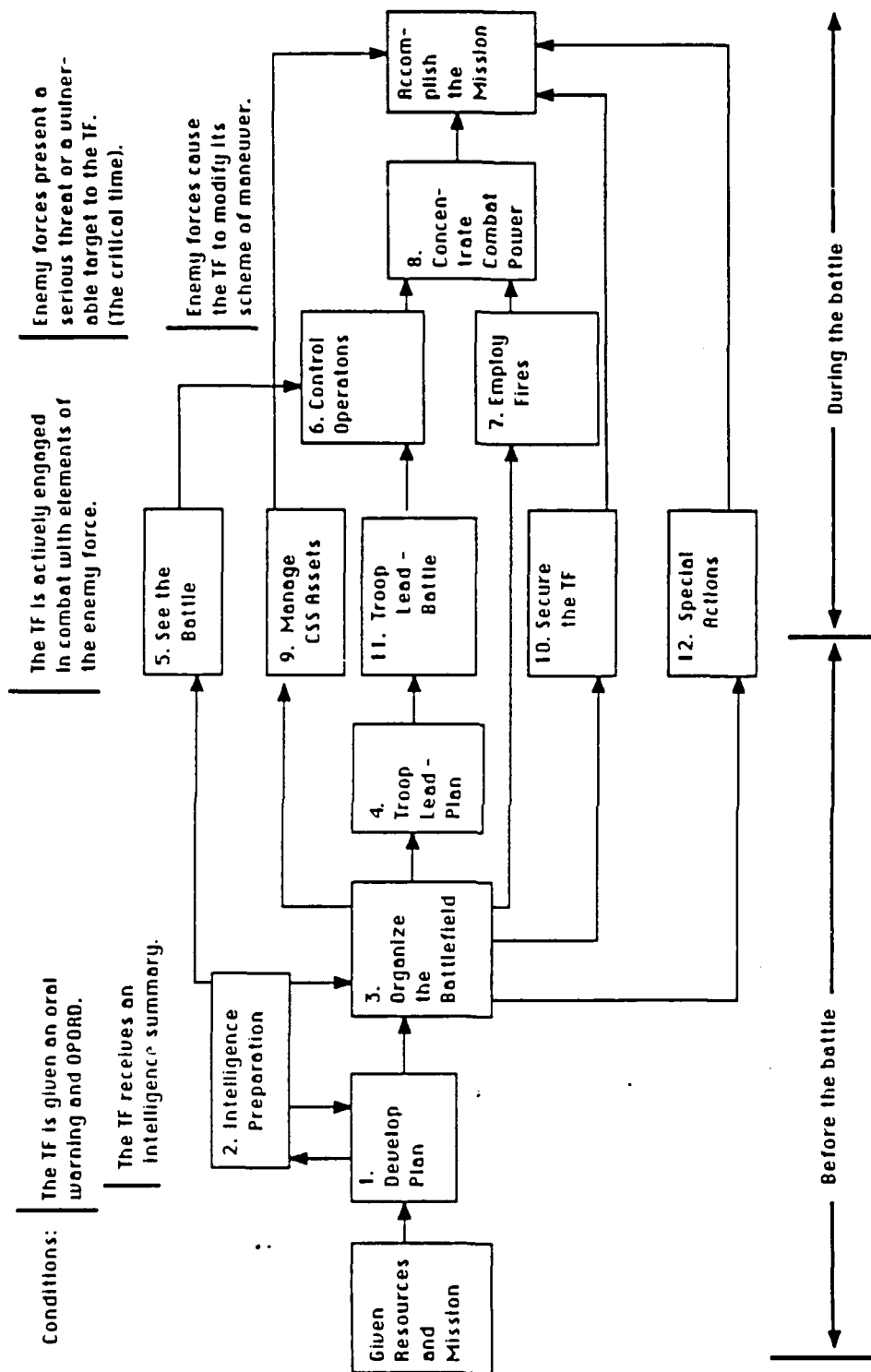


Figure 8. Flow diagram of the battalion command group ANTEP.

performance and those that were relatively weak. Figure 9 illustrates the relationships among several sets of subtasks that were identified in this study. Fifty of the 61 subtasks in the Battalion Command Group ARTEP could be evaluated in CATTs. Of these 50 subtasks, 23 were correlated at the .01 level with overall performance measures and 8 others at the .05 level. An overlapping set of 19 subtasks were rated as poorly performed. The 14 subtasks that were both poorly performed and highly correlated with overall performance are marked in Appendix H.

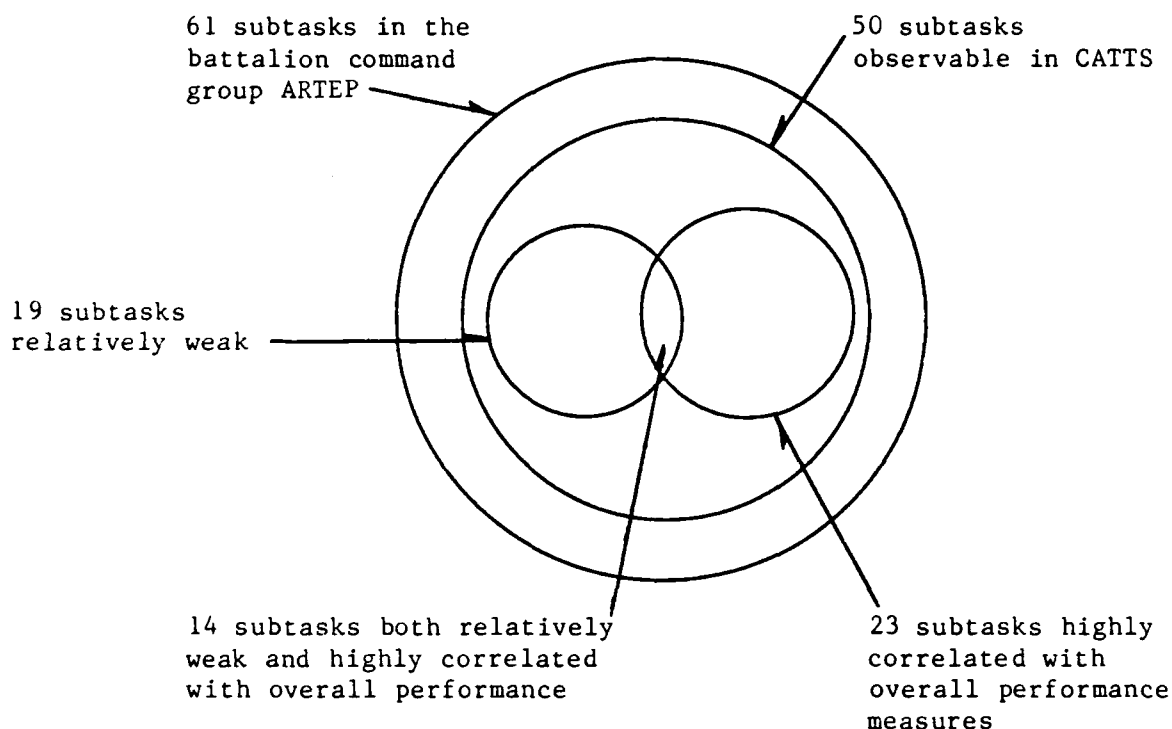


Figure 9. Summary of subtasks evaluated in CATTs.

The researchers note in the discussion of their results that CATTs does provide a suitable mechanism to conduct command staff ARTEPs, with the implied caveat, of course, that 11 of the subtasks cannot be evaluated. The authors also note that the 14 subtasks that were performed in a weak manner but were also significantly correlated with overall performance appear to constitute an essential core of the battalion command group ARTEP.

Critical ARTEP tasks. In a second CATTS based effort reported by the same researchers Kaplan and Barber (1979b) discuss research conducted:

1. To validate a revised and expanded battery of performance measures by using the measures to evaluate (a) specific command group performances and (b) the overall effectiveness of individual staff members and the command group as a whole.
2. To verify and extend the results of their previous research by using the refined assessment procedure to further identify and describe those command group performances that are most important for training in terms of (a) low performance ratings and (b) high correlations with overall effectiveness ratings.

Procedure. The research was conducted in much the same manner as the previously discussed Barber and Kaplan work. Data were collected from 23 battalion command groups that participated in Combined Arms Tactical Training Simulator (CATTS) exercises. Thirteen mechanized groups performed a covering-force mission followed by an attack, and 10 nonmechanized groups performed a defense and an attack. The groups' performances were rated on 180 items derived from the subtasks of the Battalion Command Group ARTEP. These items were rated by seven controllers who conducted the exercise and by an observer who only observed command group performance. The raters also evaluated the overall effectiveness of individual group members and of the command group as a whole.

This later research differed in the manner in which the ratings were collected. The researchers modified and improved the data collection process based on problems they had observed in the earlier research. The measurement instruments developed for this research were questionnaires answered by the experienced evaluators who observed the command group's behavior in the simulated combat environment. The questionnaire items were derived from the command group ARTEP and from previous research on the performance of battalion command groups in simulated combat (Barber & Kaplan, 1979). The previous research used a 3-point scale to rate the command groups' performance on ARTEP subtasks, but, the raters had generally avoided the low end of the scale. Therefore, a 5-point rating scale was introduced to encourage finer discriminations in performance. Also, many questions that could be answered yes or no were added to the observation form to provide specific information about certain components of subtask performance. The researchers made data collection as easy as possible by creating customized sets of forms for each type exercise (defense, mechanized and non-mechanized attack, covering force) and for each controller. There were 180 items rated and 16 tailored sets of data collection forms. Figure 10 shows a sample of the data collection sheets which illustrate the type of items and scoring. Figure 11 shows another sample with related scores and statistical data.

Results. The results generally supported the conclusions of the earlier study. Fifteen subtasks were found to be low rated on performance but to be correlated significantly with the controllers judgement of effectiveness.

Subtask	Standard and Key Events	Ratings					Comments	
4. Modify fire support plan (7A)	Priority of fires which supports the new scheme of maneuver is immediately communicated to supporting and supported units.	☐	☐	☐	☐	☐		
	Requests for immediate fire support are received and assigned to the appropriate fire support agencies. Missions are assigned which support anticipated developments.	☐	☐	☐	☐	☐		
	a. During the battle were priority of fires supporting new scheme of maneuver immediately communicated to supporting and supported units?	☐	☐	No	Yes	☐		
	b. Were Requests for immediate fire support received and assigned to appropriate fire support agencies?	☐	☐	No	Yes	☐		
	c. Overall, how well did the command group perform relative to the standards	☐	☐	1	2	3	4	5
		☐	☐	☐	☐	☐	☐	
		☐	☐	☐	☐	☐	☐	
		☐	☐	☐	☐	☐	☐	
		☐	☐	☐	☐	☐	☐	
		☐	☐	☐	☐	☐	☐	

5. Overall, how well did this command group utilize its fire support asset in comparison with previous groups in Covering Force operations?

Check one:

- ☐ 1. One of the worst. ☐ 4. Better than average.
☐ 2. Worse than average. ☐ 5. One of the best.
☐ 3. Better than average.

Figure 10. Sample of data collection sheets.

S1/S4 Ratings
Covering Force

Correlation

Item	N	%YES	MEAN	SD	S1	S4	CG
1. In providing supplies to accomplish the mission:							
a. Did the S4 coordinate with supporting supply elements?	13	69			.68*	.81**	.52
b. Did the S4 know the battalion mission prior to the LD time?	13	85			.44	.45	.09
c. Did the S4 know the Bn OPORD prior to the LD time?	13	69			.79**	.81**	.56
d. Did the S1/4 have the operational overlay in the combat trains prior to its effective time?	13	77			.75*	.64*	.52
e. How effectively did the S4 provide supplies to accomplish the mission? (3J)	13		3.15	1.46	.78**	.87**	.68*
2. In preparing for the battle:							
a. Did the S1 know the battalion's mission prior to the LD time?	13	92			.39	.21	.00
b. Did the S1 know the Bn OPORD prior to the LD time?	13	92			.39	.21	.00

*p < .01 **p < .001

Figure 11. Example of controller rating sheets, item data and analysis results.

Eight of the earlier studies 14 critical subtasks were re-identified along with seven that differed from those found in the earlier study. The subtasks that were considered critical fell into five functional areas as follows (numbers in parenthesis refer to Appendix H items):

1. Fire support: Develop (1-I) and modify (7-A) the fire support plan.
2. Intelligence preparation of the battle: Identify (1-B, 2-A), gather (2-B), and supervise compliance with the task force order (11-A). Determine the critical place and time (8-A) and concentrate/shift combat power (8-C).
3. Operations: Communicate/coordinate plans and orders (3-G) and changes (6-B), and supervise compliance with the task force order (11-A). Determine the critical place and time (8-A) and concentrate/shift combat power (8-C).
4. Logistics: Arm and fuel the systems (9-A) and integrate combat service support into the scheme of maneuver (9-D).
5. Electronic warfare: Combat enemy electromagnetic intelligence (10-A) and electronic warfare (12-A).

The authors indicate that the four observed missions were markedly different in subtask items criticality; that rater reliability was low, as low as .22 for a single rater; and that differences among ratings of the same command group by different observers were significant at beyond the .001 level. These conclusions concerning low interrater reliability and the markedly different criticality scores are matters of considerable concern. It is unlikely that these data were atypical of the kinds of data collected at ARTEPS, FTXs, and CPXs in general. Hence, there is further reason to question the value of on the spot judgements made by rater/controllers in these situations.

The researchers note also that the effects of mission type, rater reliability, and individual differences among raters have implications for the measurement of command group performance and that these effects should be controlled when diagnosing training requirements, comparing command groups, or evaluating training systems. They also conclude that the low rater reliability and significant differences among raters indicate the desirability of further research to develop more objective measures of command group performance and to identify those subtasks where the difference in perspective of raters might cause valid differences in performance ratings of those raters.

CATTS as Surrogate ARTBASS

By the end of the decade of the 70s the Combined Arms Center, who had responsibility for development of command and control training systems, had accumulated considerable experience with both the CATTS and the CAMMS systems.

Based on data obtained during the use and refinement of these systems a decision had been made to develop a follow on system, the Army Training Battle Simulation System (ARTBASS). The system was considered as an evolutionary development from CATTs which at the time was considered the most advanced battalion command and control training vehicle and research test bed.

Both of the research efforts discussed here were motivated by a requirement to demonstrate that ARTBASS would indeed be effective for training battalion command groups. Due to the compressed developmental schedule of ARTBASS, and the close similarity of ARTBASS to CATTs, the ARTBASS Test Integration Working Group decided that a Training Development Study (TDS) could be conducted on CATTs as surrogate for ARTBASS. A three phase research effort was proposed. Phase 1 was to be designed to identify system and scenario characteristics which effect group performance in simulated battles. Phase 2 was to address behavioral variables that influence battalion command group effectiveness and identify training strategies to increase training benefits. Phase 3 was to be an attempt to combine knowledge gained during the two earlier phases to develop an optimal set of training exercises for battalion command groups. The research projects discussed in this section were designed to address the first and second goals. Apparently the CATTs phase out occurred so rapidly that the third effort, which could not be planned until data from both earlier studies were analyzed, could not be conducted.

Surrogate research 1. The primary focus of the first research effort (Thomas, Barber, and Kaplan, 1984) was to develop measures of command and control performance that were appropriate within the context of CATTs training exercises. Appropriate measures should be responsive to systematic manipulations in exercise difficulty and to training benefits derived by staff groups as a function of repeated exposures to CATTs training. To this end, the research was designed to determine if, and to what degree, a selected set of simulated battle variables could affect the difficulty of CATTs exercises as assessed by battle outcomes, ratings of command group performance, and ratings of exercise difficulty. Ratings of exercise realism were also collected from each group of raters as a check on the fidelity of the CATTs simulation. Data were obtained during the period October through December 1981. Five battalion command groups participated. Each group was tested in a sequence of four, one-day CATTs exercises. The system and scenario characteristics varied were terrain, mission, weather, communication, and combat ratio. The levels for each were as follows:

- | | |
|-------------------|---|
| 1. Weather played | Clear visibility vs fog or dust |
| 2. Terrain | Fulda Gap (hilly) vs Sinai (level) |
| 3. Communication | Perfect communication with token jamming vs increased jamming and less land wire. |
| 4. Mission | Attack vs covering force |
| 5. Combat ratio | Own 100% / OPFOR 75% vs Own 80% / OPFOR 100% |

The researchers used a five-factor fractional factorial design (see Daniel, 1976) which permitted them to use a total of 16 combinations rather than the 32 which would have been required for a full factorial design. Since only four

battalions were required to provide the 16 combinations the fifth battalion command group tested was considered a back-up group whose data could be used to fill in if computer problems caused loss of computer generated data. This loss occurred on the first day's exercise of the second battalion command group cycle so the missing combinations were played for the day one exercise of the back up group.

Three types of data were collected during the research event; performance measures, system measures and information flow measures. The information flow portion of this research and of the following Thomas, Kaplan and Barber, 1984, research were discussed previously in the section entitled, "Who Knew What, When" and will not be discussed again here. The simulation outcome data for both studies are discussed in a later section, "Command Post Effectiveness, What to Measure." These data also will be discussed only to the extent needed to describe the research results relevant to this section.

The researchers provided analyses of results obtained from subjective ratings provided by controllers, players, and player/controllers. The controllers provided subjective ratings on up to 27 ARTEP derived tasks, an overall rating of how well the command group performed, specific ratings for the BC, S1, S2, S3, FSO, ALO and each of the company commanders; a difficulty score for each of 14 planning/execution items; and a realism assessment score for each exercise.

Inter-item correlations for the 27 ARTEP derived items which had been developed and used in earlier ARI research (see Barber & Kaplan, 1979; Kaplan & Barber, 1979b; and Barber & Solick, 1980), for the 14 items concerning difficulty and for those ARTEP and difficulty items that "correspond" are shown in Table 17. Intercorrelations were highly significant for all groups even though players and player/controllers (the company commanders) median correlations were quite a bit lower. The most interesting results were obtained when the researchers performed factor analyses on each of the six data sets: controller, player, and player/controller data for the 27 performance items and for the 14 difficulty items.

Only one factor was derived from the 27 ARTEP controller ratings. This suggests that the controllers scored the 27 items high or low according to some general feeling that effected all the items rather than on factors relevant specifically to the individual items themselves. Recall at this point that in a previously discussed study controllers indicated that they did not feel they had time to be good observers. Also, that in another study the reviewer felt compelled to point out that since the controllers knew the outcome (and in this case at least the progress) of the battle before they made their ratings of performance there might be considerable positive halo effect. The presence of only a single factor in the analysis of the 27 ARTEP based performance measures clearly supports the conclusion that controllers are not able to perform the required ratings in a valid way even though they seem to agree significantly on their ratings. Halo effect - you won so you must have performed well - appears even more attractive an explanation than when it was advanced earlier.

Table 17

Inter-Item Correlations of Performance and Difficulty Items

		Performance Items (27)	Difficulty Items (14)	Between Corresponding Performance and Difficulty Items
Controllers	Median	.88	.85	.67
	Range	.33 - .98	.53 - .98	.43 - .89
	N*	41 - 124	85 - 128	78 - 126
Players	Median	.54	.51	.34
	Range	.35 - .95	.27 - .83	.11 - .54
	N*	144 - 191	117 - 159	117 - 158
Player- Controllers	Median	.62	.61	.54
	Range	.37 - .92	.35 - .87	.36 - .74
	N*	101 - 139	78 - 115	76 - 107

* Difference in N-sizes were due to the fact that not all raters rated all items.

Players and player/controllers were better able to evaluate specific task performances. Player data provided three factors. Factor I was equivalent to Task 1 of the ARTEP performance standards (see Appendix I). "Gather and analyze required information." Factor II was nearly equivalent to the combined tasks 2 and 3: "Develop a plan based on mission and modify it as required by events," and "Communicate/coordinate". A third factor was roughly equivalent to tasks 4 and 5 "Implement plan" and "Supervise combat operations." The factor analysis of the performance ratings recorded by player/controllers resulted in two factors: Factor I representing tasks 4 and 5 and four subtask items from task 2; and Factor II was equivalent to tasks 1 and 3 plus eight subtasks from task 2.

The factor analysis performed on exercise difficulty questions indicated planning difficulty and execution difficulty factors for both the controllers and player/controllers. That is, Factor I included all items referring to planning, and Factor II included all items assessing the difficulty of execution. The analysis of player ratings of exercise difficulty also resulted in two factors. Factor I included all items except those assessed in the difficulty of planning and execution of commo and admin/log. Items not in Factor I were incorporated in Factor II.

The small number of factors and the relatively high inter-item correlations suggest that the raters had difficulty discriminating among the ARTEP items created to assess battalion command group performance and items related to exercise difficulty. It appears, however, that the players and player controllers were better able to make these discriminations than were the controllers.

Data from performance and difficulty ratings by controllers, players, and player/controllers were also analyzed across the four exercise days. No note-worthy conclusions resulted from these analyses although generally raters seem to feel that exercises became more difficult and performance became better over the four days of exercises. The authors note that four exercises by four groups - 16 group exercise combinations - is not a large sample and that even though some of their statistical comparisons were significant interpretation is difficult. It is interesting to note that there were two sets of measures that were objective rather than subjective: battle outcome and information flow. When battle outcome data - RER, SMFRD, C²ILL and CR were analyzed by using a series of 2x2 analysis of variance models (attack and covering force missions and first versus second battles) none of the results were significant. When the communication data, obtained by a carefully controlled insertion of information and a test of individual knowledge, were analyzed it was found that communication did not increase significantly. The conclusion that the experimental data were too few is immeasurably more attractive than the alternative explanation - that CATTS the ARTBASS surrogate does not teach anything.

Analysis of the terrain, mission, weather, communication and combat ratio variables in terms of the four combat outcome variables gave encouraging results. A summary of the results of the four analysis of variance are shown in Table 18. Probably the only surprising result is the fact that communication was not a significant variable. Terrain certainly effects both the attacker

and the defender and, good or bad, it may even out. But poor communications - which appeared to have been a one sided problem - should degrade the possessor. Perhaps the critical degradation point had not been reached and command groups had been able to compensate for non-significant degradation.

Table 18

Summary of Statistical Significance Level of Five Main Variables for the Four Simulation Outcome Measures

Variable	Simulation Measures			
	RER	SMFRD	CR	C ² ILL
Terrain	NS	NS	NS	NS
Mission	<.005	NS	<.025	< .05
Weather	< .05	< .05	NS	.05
Communications	NS	NS	NS	NS
Combat Ratio	<.005	<.005	<.005	<.005

Surrogate research 2. In the second study that evaluated ARTBASS using CATTIS as a surrogate Thomas, Kaplan, and Barber (1984) proceeded as follows. Five battalion command groups (players), from three mechanized infantry and two armor battalions, participated in four one-day CATTIS exercises between December 1982 and March 1983. The first and last exercises were designated as pretest and post-test exercises, respectively, and included covering force/delay missions on different portions of Fulda Gap terrain. The intervening training exercises were delay and movement to contact exercises conducted on either on Ft. Irwin or Sinai terrain. Measures of performance included (1) information reception and transmission by group members during planning (information flow questionnaire), (2) information exchange by staff members during battle execution (probes), (3) degree of success on the simulated battlefield (mission accomplishment scores), and (4) ARTEP performance as assessed by CATTIS controllers, player/controllers, and the players themselves. The data collection forms and processes were the same as in the Thomas, Barber, and Kaplan (1984) research just described.

This research differed from the earlier research because of the inclusion of a feedback mechanism that provided players specific indications of their performance and gave them an opportunity to discuss potential corrections. In the battle replay sessions at the end of each exercise day, CATTIS controllers

replayed significant events on a video display of the terrain, which showed the position and strength of friendly and enemy forces during the battle. Then the players and player-controllers presented their perceptions of the battle in after-action reviews. Feedback conferences were held on the morning following the pretest exercise. ARI personnel provided the battalion commander with results from the information flow questionnaire and with ratings of staff performance on a short list of ARTEP tasks. Problem areas were identified and solutions were recommended. The other staff members (S1, S2, S3, S4, FSO, and ALO) met with their brigade controller counterparts and discussed problems identified in the pretest exercise. Discussions included, but were not confined to, performance on ARTEP tasks and the group's ability to transmit important information during mission execution. Problem areas were identified and tentative solutions recommended by controllers. At the conclusion of the third exercise day, feedback conferences were again conducted in the same way and with the same people. Discussions focused on problems that had been resolved, new problem areas, and recommended solutions to problems.

Two of the sets of measures, information flow and mission accomplishment, have been discussed elsewhere at some length so they will be only briefly treated here. Given feedback and help in defining potential corrective actions the amount of recalled information increased at each of the communication stages. In two cases - within the staff group and by company commanders - the increase was greater than in the Surrogate 1 research. In the other case - brigade to battalion - the increase was the same but in this research it was statistically significant. (Possibly due to more cases since in this project all five battalion scores were used rather than the four used in the previous analysis, or, perhaps, to a chance reduction in variance, or both.)

The researchers used a mission accomplishment measure developed by Thomas and Cocklin (1983). The measure, which was applied to pre and post-test data, requires judgements from experts concerning the success implied by the after the battle state of five factors. The researchers reported that an analysis of variance applied to mission accomplishment scores showed significant differences between units ($p < .05$), between first and fourth day performances ($p < .05$), and a significant unit by day interaction ($p < .001$). The first difference is not unexpected. The third is evidence that although training with feedback improves performance some groups change more than others. For example, although four of the five battalions had improved post-test scores the fifth did much worse on post-test.

The researchers did not factor analyze their ARTEP task based data so it is not known if the same factor relationships as those discussed in the first CATTS surrogate study again existed. It seems likely, however, that the problem of controller data all being based on one factor and others on only a few also existed since the median inter-item correlations for controller, player, player/controller data were quite similar: .85, .51, and .61 in the previous study and .80, .71, and .65 in the current study. Because of this apparent lack of discrimination among items the researchers used an average computed

from total daily ARTEP scores for their later analysis. Generally, the researchers found that ARTEP item scores for the controller, player, and player/controller ratings were all significantly higher on the post-test.

This research complements the initial surrogate research quite well. Clearly, performance feedback and discussion of how to improve performance makes the simulation based training exercises more effective battle management training experiences. Equally apparent is the conclusion that the use of controllers, players, and player/controllers themselves to provide ratings based on a wide variety of ARTEP like tasks is unwise since it leads to much data but to data that have but little meaning.

BRIGADE AND DIVISION COMMAND AND CONTROL RESEARCH

The amount of research that has been conducted at echelons above battalion is quite limited. Indeed the brigade material appears to be limited to four reports and the division material to four also. Some additional work has been done at higher level command posts but since it dealt with naval command posts at what might be described as echelons-above-corps its relationship to Army command and control processes is limited. This latter work, which deals with a command and control measurement technique called Headquarters Effectiveness Assessment Tool (HEAT), will be discussed in a later section.

BRIGADE LEVEL RESEARCH

One of the brigade level research events has been discussed earlier in the section "CAMMS versus a CPX." In this study, which considered both battalion and brigade command groups, the researchers investigated how effective and cost efficient CAMMS based ARTEPs were when compared to CPXs. To obtain data questionnaires were administered to 50 battalion command groups and 12 brigade command groups after they participated in CAMMS exercises. The questionnaires asked players to rate CAMMS and a CPX on several measures of training effectiveness, including realism, motivation, and the degree to which the player was required to perform ARTEP subtasks. Estimates of man-hours required to prepare CAMMS and a CPX and of the number of controllers needed were provided by 14 CAMMS exercise directors and for the CPX, from 5 division and brigade commanders.

The researchers reported that CAMMS exercises were judged significantly and consistently more realistic and more interesting than a CPX. Two functional areas in which CAMMS enjoyed the greatest advantage were related to preparing and organizing the battlefield, and controlling and coordinating combat operations then concentrating combat power. Commanders felt CAMMS did a good job of exercising command and control processes, especially in exposure to the capabilities of enemy weapons systems, in facing a thinking enemy, and in making decisions under realtime constraints. CAMMS exercises also cost 25% to 30% less than a CPX, primarily because CAMMS required much less preparation time.

Principal weaknesses of CAMMS were that it did not produce much stress and did not exercise the players in security procedures, such as electromagnetic and communications security. Nor did CAMMS require players to react to special situations, such as enemy jamming or chemical, biological, or nuclear warfare. The CPX also received low ratings in these areas. Despite these shared weaknesses the researchers concluded that overall, CAMMS produced a distinctly superior exercise at a moderate savings in cost over a conventional CPX.

The remaining brigade based work was conducted by Olmstead and fellow HumRRO researchers. The project, which led to three reports relevant to this review, was designed to examine the contribution of organizational processes to brigade effectiveness in CAMMS based exercises and to determine the effects of brigade command group supervision and control on command group performance of

those processes. The project also sought to develop improved strategies and methods for training brigade command groups. The authors described their project objectives as:

1. Using CAMMS performance as the source of data, to determine the relative criticality of brigade command group ARTEP tasks and relationships among the various ARTEP subtasks.
2. To develop recommendations for ARTEP modification.
3. To develop training objectives for the brigade command groups functioning as a whole.
4. To develop average ARTEP performance data for brigade command group participation in CAMMS.
5. To develop a CAMMS training feedback system for use with brigade command groups.
6. To study and identify factors or patterns of organizational behavior within brigade command groups which contribute to more or less effective unit performance.

The three reports that resulted from this research dealt with aspects of the problem of training and measuring the performance of brigade command groups. The first report (Olmstead, Baranick, & Elder, 1978a) addressed the relationship between ARTEP task ratings and CAMMS battle outcome scores. Thus, it covered the first five of the six objectives listed above. A second report, Olmstead, Baranick, & Elder, 1978b, considers how to best use CAMMS, and by implication other simulations, in a command group training process. The third report Olmstead, Baranick, & Elder, 1978c) addressed the relationships between organizational processes, organizational effectiveness, and command group supervision and control.

All of the reports were based on the same data which were collected on 11 brigades, 8 active Army brigades, 2 Reserve brigades, and one National Guard brigade. Researchers collected data as the units' participated in CAMMS based CPXs. Three types of data were collected. One type of data was needed to support analyses of ARTEP task performance. The other two data sets were required to permit a replication and extension of the earlier Olmstead, Christensen, and Lackey (1973) work that had investigated the relationship between organizational competence and organizational effectiveness in battalion command posts. (See section titled "In the Beginning. . ."). To accomplish the first goals the researchers collected the following types of data:

1. Battle outcome results of CAMMS exercises.
2. Ratings by division-level controllers of overall brigade command group effectiveness, based upon controller observations of command group performance of tasks specified in draft ARTEP 100-1 for brigade command groups.

3. Ratings by three levels of players (brigade, battalion, and company) of the quality of organizational processes within the brigades and of a number of dimensions subsumed under the rubric supervision and control.

The researchers also obtained measures of the following nine process dimensions:

1. Information acquisition
2. Providing information and intelligence
3. Anticipating contingencies
4. Timeliness of adjustments in plans and operations
5. Effectiveness of adjustments in plans and operations
6. Planning
7. Decision making
8. Coordination
9. Communication

And the following six supervision and control dimensions:

1. Clarity of mission objectives
2. Clarity of roles
3. Responsiveness to subordinate unit requirements
4. Quality of supervision
5. Amount of supervision and control
6. Delegation

In addressing the issue of how ARTEP task ratings relate to battle outcome results it is clear that the researchers had the same problems that have already been discussed in reviewing battalion level research. Some of the difficulties were related to the nature of ARTEPs and ARTEP tasks. The 12 ARTEP tasks were considered too broad to serve as a basis for correlational analyses with battle outcome data, hence, subtask scores were not calculated. This in turn led to the problem, already discussed elsewhere, that not all ARTEP subtasks were played either because CAMMS could not play them, the scenario could not support them or the exercise planners simply opted not to play them. In all, data were collected on only 42 of 61 ARTEP subtasks. Battle outcome was quantified by using the area controlled by resources remaining model described by Tiede and Leake (1971).

In analyzing their data the researchers computed correlations (Pearson r) based on three sets of measures; ARTEP subtasks, the mean of the controllers ratings on "overall effectiveness" of each command group, and mission accomplishment scores based on the area/resources model. The researchers reported that all ARTEP subtasks were positively correlated with command group effectiveness, with many of the items having moderate to high relationships and that eight subtasks had correlation coefficients significant at or beyond the .05 level. The authors felt that it was meaningful that 29 subtasks showed correlation coefficients of .40 or higher, the point which they noted, is conventionally taken as indicating a substantial relationship. The 29 substantially related subtasks were divided approximately equally between activities performed before the battle (14) and those performed during the battle (15).

The relationship between ARTEP subtask scores and the battle outcomes was far less encouraging. Of the 42 subtask scores, 17 were positive, not statistically significant, and indeed, as a group low enough to appear insignificant (.43 to .02). On the other hand 25 of the subtasks were negatively correlated with battle outcome and, of these, four were high enough to be statistically significant. One could conclude that the relationship between a single ARTEP subtask and the outcome of a brigade battle engagement is too tenuous to create the expectation that it should be significantly related. On the other hand a reader could look at the tilt toward an inverse relationship and conclude that good performance on ARTEP tasks increases the likelihood of defeat. Clearly the former conclusion is more attractive. The authors in this study conclude that it was apparent that no meaningful relationships existed between ARTEP subtasks and their battle outcome measure. They hypothesize, however, that it may be because of "the highly questionable validity of the combat results data" and they state that "the fact that no relationships were found does not disprove the validity of any ARTEP subtask."

ARTEP subtasks scores were generally related. Of 861 correlations, 130 were significant. Numerous other moderate, but not significant, relationships were found. For the most part, the strongest relationships follow logic. That is, most significant relationships are grouped among subtasks which fall within the same activity areas, such as planning, fire support, etc., or which could logically be expected to be closely related, e.g., several planning subtasks and organizing for combat. However, the authors feel that the fact that many of the subtasks were not related indicated that they probably represented discrete command group activities, each of which might contribute independently to unit effectiveness if their validity could be proven.

The authors do not provide an extended discussion of the relationship between ARTEP tasks, their overall command group effectiveness score and their battle outcome score. However, examination of the tabular data of the 42 ARTEP subtasks and their relationship to the overall command group effectiveness and battle outcome scores shows the following range and medians for the 42 correlations: ARTEP subtasks to battle outcome, range .43 to -.72 and median -.13, ARTEP subtasks to overall effectiveness, range .99 to .05 and median .48. It is clear that the results of this study supported other work by the same research team (See Olmstead, Christensen, and Lackey 1973 and Olmstead, Elder,

and Forsyth, 1978c) which demonstrated that subjectively developed ARTEP measures do not correlate well with equally subjective command group effectiveness measures and correlate even less with battle outcome scores.

Since the basic purpose of this research project was the development of improved command group training the researchers defined 11 terminal learning objectives and 33 associated learning objectives for brigade command groups. They also recommended changes in some ARTEP tasks and standards to ease the evaluators difficulties in scoring the subtasks. The reader who is interested in the ARTEP subtasks that were found usable in the CAMMS simulation environment will find the list in Appendix J. Appendix K contains a list of terminal and associated learning objectives.

The researchers provided two more reports which were based on the data obtained in the research described above. In the first of these reports (Olmstead, Baranick, & Elder, 1978b) describe a training feedback system for brigade command groups which involve CAMMS based exercises. The report summarizes information from experiences of the researchers as they assisted in the brigade command group exercises.

The final report provided by the researchers (Olmstead, Baranick, & Elder, 1978c) contains the most interesting analysis of their data. This report provides the analyses needed to support the sixth objective on the list shown earlier. The objective was, of course, of such a nature that it afforded the researchers an opportunity to collect data to further investigate the relationship of organizational competence to organizational effectiveness. In a research project discussed earlier (Olmstead, Christensen, & Lackey, 1973) Olmstead and his associates had demonstrated that there was a very high relationship between how competently - how well - organization personnel performed the functions required to cope with a changing environment and the command groups effectiveness, as measured by the quality of its decisions. Later work, analyzed and reported almost parallel with this brigade research, showed that the relationship was still significant but much lower when ARTEP tasks and battle outcome scores replaced the more rigorous and more direct measures of competence and effectiveness used in the 1973 research. (Olmstead, Elder, & Forsythe, 1978c).

In this research the researchers modified their conceptual framework, or model, based on the experiences of their earlier research efforts and extended their model to more specifically address factors relating to the supervision and control of subordinate units. This extension helps the model reflect the often noted dichotomy between battle planning activities and the processes which are involved in fighting the battle. The following organizational processes and associated dimensions were defined and used as the basic framework for analysis:

<u>Process</u>	<u>Dimension</u>
Information Acquisition and Processing	Information Acquisition Providing Information and Intelligence
Adaptability	Anticipating Contingencies Timeliness of Adjustments in Plans and Operations Effectiveness of Adjustments in Plans and Operations
Implementation	Planning Decision Making Coordination
Communication	Communication
Supervision and Control	Responsiveness to Subordinate Unit Requirements Amount of Control and Supervision Delegation Clarity of Objectives Clarity of Roles

The main thrusts of the study were to examine the relationships between the various organizational processes and unit effectiveness, and to evaluate the effects of supervision and control of subordinate units upon performance of the other four processes. In discussing their research the authors noted that the effective performance of the processes involved in competence is one of the principal ways that a command group influences the effectiveness of subordinate units and, thereby, the outcomes of military operations. The effects exerted upon a combat operation by the command group of a large unit such as a brigade are the results of the acquisition, processing, and interpretation of information; the development of operational concepts and plans; the diffusion of plans, orders, and instructions through successively lower level units; and the supervision of implementation of the plans by lower level units. Thus, the fundamental military function, known as command and control, involves mainly the execution of competence processes that result in effective decisions or plans, and supervision and control of their execution by subordinate units. The authors stated that, when brigade command groups perform the command processes effectively and provide high-quality supervision and control to subordinate levels, equally high-quality process performance should result at lower levels. Thus, if process performance, and supervision and control are both better, unit effectiveness should also be better.

Data collected at the 11 brigade CPX events consisted of controller evaluations, player evaluations, and battle outcome results. Controllers rated brigade command group performance on 42 ARTEP subtasks, on five activities defined to represent organizational processes, and on five "overall" measures. The

five organizational process measures were: planning, decision making, implementation, communication, and responsiveness to subordinate units. The five overall measures were: total command group, S1, S2, S3, and S4. The 42 ARTEP standards, other measures and scales used information are shown in Appendix I. The authors note that the quality and validity of the data obtained on controller rating forms and the combat results data were of questionable quality but that the player evaluations were probably valid since the form was designed to obtain player perceptions. This form is provided in Appendix L.

Table 19 shows a summary of the results obtained when the combat effectiveness (battle outcome) and the overall command group effectiveness scores were correlated with the nine process dimension scores. The reader will note that none of the process dimensions were significantly correlated to the area controlled - resources remaining based combat effectiveness measures. Indeed the highest correlation was a $-.54$ between combat effectiveness and the planning

Table 19

Correlation of Process Dimensions with Brigade Combat Effectiveness and Command Group Effectiveness

Dimension	Combat Effectiveness		Command Group Effectiveness	
	r	p ^a	r	p ^a
Information acquisition	.24	NS	.81	<.002
Providing information and intelligence	.44	NS	.79	<.004
Anticipating contingencies	.03	NS	.67	<.02
Timeliness of adjustments in planned operations	.06	NS	.82	<.002
Effectiveness of adjustments in planned operations	.02	NS	.74	<.01
Planning	-.54	NS	.16	NS
Decision Making	.04	NS	.83	<.002
Coordination	.33	NS	.02	NS
Communication	.35	NS	.72	<.01

a. Degrees of freedom = 9.

process dimension. The relationship between the process dimensions and the general combat group effectiveness rating, were highly significant for seven of the nine items. Of course the same subjective rating process was applied to obtain the process and overall rating, and the same raters were involved so it is likely that the relationship was inflated somewhat because of rater bias and halo effect.

Table 20 shows the results of analyses conducted after dividing the brigade into two groups: the five most combat effective and six least combat effective. Four of the nine process dimensions showed significant differences. It appears that command groups who anticipate contingencies and make better decisions that lead to timely, effective adjustments do better than those who do not do these things as well.

Table 20

Process Performance of Brigades with Most and Least Effective Command Groups

Process Dimension	Process Performance				t	p
	Most Effective Mean	Least Effective SD	Least Effective Mean	SD		
Information acquisition	4.51	.15	4.13	.51	1.60	NS
Providing information and intelligence	4.52	.14	4.20	.57	1.22	NS
Anticipating contingencies	5.25	.24	4.76	.34	2.73	<.05
Timeliness of adjustments in planned operations	5.40	.18	4.77	.40	3.25	<.01
Effectiveness of adjustments in planned operations	5.59	.16	5.03	.45	2.66	<.05
Planning	4.90	.27	4.48	.45	1.81	NS
Decision making	5.39	.27	5.01	.27	2.33	<.05
Coordination	4.40	.47	4.55	.25	.68	NS
Communication	5.10	.26	4.60	.66	1.59	NS

Data were also collected from the brigades on the six previously mentioned supervision and control dimensions. The results of analyses performed on these data are shown in Tables 21 and 22. Note that no dimensions correlated significantly with combat outcome but that four of six correlated significantly with the overall command group effectiveness measure. It would appear from the data in Table 21 that good delegating behavior and clearly defined objectives result in better battle outcomes when they are supported by more and better supervision and control. Table 22 which shows data from good versus less able units generally supports this conclusion. The results can be summarized by stating that clear objectives, supported by more and better supervision, and good higher headquarters responsiveness to subordinate unit requirements

Table 21

Correlations of Supervision and Control Dimensions with Brigade and Command Group Effectiveness

Dimension	<u>Brigade Effectiveness</u>		<u>Command Group Effectiveness</u>	
	r	p ²	r	p ^a
Clarity of objectives	.19	NS	.73	<.01
Clarity of roles	.13	NS	.42	NS
Responsiveness to subordinate unit requirements	-.03	NS	.59	NS
Amount of supervision and control	.02	NS	.70	<.02
Quality of supervision	-.16	NS	.79	<.01
Delegation	.26	NS	.70	<.02

Table 22

Supervision and Control in Brigades with Most and Least Effective Command Groups

	Player Ratings ^a				t	p
	Most Effective		Least Effective			
	Mean	SD	Mean	SD		
Clarity of objectives	6.06	.29	5.42	.62	2.12	<.05
Clarity of roles	6.08	.35	5.76	.25	1.81	NS
Responsiveness to subordinate unit requirements	5.44	.35	4.76	.60	2.21	<.05
Amount of supervision and control	6.22	.20	5.75	.34	2.69	<.025
Quality of supervision	5.71	.11	4.88	.49	3.64	<.005
Delegation	6.23	.13	5.97	.42	1.31	NS

^a Most Effective n = 5; Least Effective n = 6

enable the unit as a whole to fight a better battle. These data are also subject to the caveats previously stated. Despite these limitations it does appear that the research confirms the concept that improved performance on command processes and improved supervision and control does improve overall command group effectiveness.

Certainly this report, and other reports of this research team do lead to the conclusion that improvement in ARTEP task performance, improvement in process dimensions and/or improvements in supervision and control of subordinate unit actions do improve the overall effectiveness of command groups. It is equally apparent, however, that no relationship to battle outcome was demonstrated.

DIVISION LEVEL RESEARCH

Research dealing with measuring the performance of division level command and control is virtually non-existent. Mention has been made several times in the course of this review to a study by Tiede and Leake (1971) in which area controlled and resources remaining were combined with expert military judgment to order a set of area controlled x resources remaining combinations from best to worst. The rank of a particular combination could then serve as a unit combat outcome effectiveness score. The measure was developed to support an early study conducted to determine the contribution of an improved tactical information system to the combat effectiveness of a force.

The study is of interest for several reasons. Obviously it described a battle outcome measure that had an impact on later research. Also there is an historical interest since the work was done in the late 1960s and, indeed, was first reported at the November 1969 meeting of the Military Operations Society in San Diego, California. At which time it won a prize as best paper. It is also noteworthy that the problem addressed by the paper was how to measure the contribution of a computerized tactical information system on combat effectiveness not on command and control effectiveness. Clearly, at least in the minds of the researchers there was no need to attempt to measure command and control since the project dealt directly with force effectiveness.

Much of the material in this paper deals with aspects of the command and control performance measurement problem that are only of marginal interest in this review, e.g., why ADVICE II was selected as the division combat simulation and the combat effectiveness measure development process. What is of interest is that the research was an attempt to measure the performance of division level command groups with and without a computer based system designed to improve the flow of tactical information. Despite the prize winning nature of the paper it is extraordinarily niggardly in its presentation of information concerning what was done and what data were available from the experiment. It appears, after carefully reading, that eight division battle simulations were run and that six of the events provided data that could be used to determine if the tactical operating system was useful.

Results, according to the authors, were quite surprising: the system improved the processors reaction time from a range of 3.75 to 5.5 hours to a range of 2.25 to 2.75 hours (reaction time was undefined), and enabled a defender to hold more area while losing fewer people than was possible without the system. Overall, this paper somehow manages to address tactical decision making without ever really addressing decision makers or the decision process. It is difficult to understand why the area controlled/resources remaining based battle outcome measure was so well accepted considering the nature of its base.

Division level command and control research improved only slightly in subsequent years. Clearly the concept of doing research in an actual CPX was a problem so apparently no such research was conducted. Division CPXs are not inexpensive, inconsequential events and apparently no research group was able to identify research that could be conducted during the regularly scheduled FORSCOM training. To counteract this situation attempts were made to create a simulation facility in which portions of the division command and control process could be researched. Two of the studies conducted in the Simulated Tactical Operating System (SINTOS),¹ an early attempt to create a dedicated command and control research facility are worth noting. The earliest of these studies (Krumm, Robins, and Ryan, 1973) was an attempt to develop a reliable measure of decision quality so that decision quality measures could be used, in SINTOS, to assess the influence of various parts of information systems on the quality of command and control decisions.

In the Krumm, et al. (1973) research only the G-3 planning function was tested. Thus, the research involved determining how one person made selected command and control decisions rather than how command groups performed. A total of 20 lieutenant colonels and colonels participated in the experiment. All the officers were on active duty status in the greater Washington area. Their ages ranged from 35 to 49 years and their military service time from 13 to 28 years. All but one were college graduates and eight had had some graduate school training. All subjects had had combat experience, three in World War II, ten in Korea, and eighteen in Vietnam. Total combat experience ranged from 9 to 37 months. All were infantry officers whose training and experience in the mechanized infantry prepared them for the Assistant G-3 Plans role they filled in the experiment situation.

The experiment proceeded in the following manner. The subjects were placed in the test situation and provided a copy of selected portions of a corps OPORD which assigned them a mission. In the test situation the subject could request, by phone, any information he desired and the experimenter would provide it on a CRT display at the subject station in the simulated, high-tech, plans cell. If the subject requested it hard copy printed information was provided by a printer in the subjects area and maps, or other such information, were delivered by messenger.

¹SINTOS was an effort by the Army's Behavior and System Research Laboratory (BESRL) to develop a dedicated facility. When BESRL became a part of ARI SINTOS was continued as an ARI facility.

Each subject saw the same OPORD and had the same information available. Each subject was scored on three subtests. During Subtest I, the subject was requested to identify enemy avenues of approach into his division sector, to identify key terrain features, and to recommend a form of defense. It was recognized that identification of avenues of approach and key terrain is normally a G2 responsibility, but military experts had indicated that these items could be evaluated by subjects with G-3 or commander experience.

When the subject had completed Subtest I, he received "commander's guidance" which indicated that the "commander" had elected to use an area defense and, as Subtest II, wanted the subject to develop a course of action based on that concept. The subject was requested to allocate maneuver elements to the echelons of defense (including a task force on the general outpost line, his forward defense area forces, and his reserve forces) and to recommend the type of resistance to be offered by each echelon of defense. The subject recorded his decisions on data sheets provided.

At the completion of Subtest II, the subject again received "commander's guidance" regarding the course of action, and, as Subtest III, was requested to develop the graphic portion of the defense order, sketching it on his situation map overlay and indicating his recommended location of the general outpost, combat outpost coordination point, brigade boundaries, forward defense area, battalion positions, reserve force location, anticipated allowable enemy penetrations, and blocking positions. The subject then wrote his recommended task organization and mission statements to subordinate units.

An experimenter log was maintained which listed the requests for data made by each subject, including the essential content of the request, time of the request, and time at which the response was sent. The record included a notation of whether the data were displayed on the CRT or on the projection screen and whether hard copies had been requested. After the test, each subject completed a questionnaire concerning his military experience, and his civilian and military education.

Data collected from the subjects and during the experiment were then used to create four measures which essentially described the subjects experience, ability, decision process pattern and facts he had available as he completed each subtask. The material used for the experiment was developed from Command and General Staff College course material. Hence, there was a "school house" solution for each of the subtest tasks. The school house solution therefore served as criteria for scoring the subjects responses. A "consensus standard" scoring technique was also developed by which individual responses were compared to what the authors refer to as "group average responses."

Data were analyzed and the relationship between four predictor variables and the criterion variables were determined. The researchers report a multiple correlation of .91 between the prediction variable data and the decision quality criteria based on the school house solution. (The consensus score was found inadequate and dropped from consideration). The researchers concluded that their work had shown the feasibility of using schoolhouse solutions as

criteria, and their scoring algorithm, shown in Appendix M, makes a contribution to the command post effectiveness measurement problem - albeit a fairly complex one.

A second study in the SIMTOS environment (Robins, Buffardi, and Ryan, 1974) was much the same as the earlier one. The SIMTOS equipment had been modified by the time this research was undertaken so it was possible for subjects to obtain CRT data without calling the experimenters to request it. The scenario was a CGSC developed division-in-defense scenario. The subjects, acting as an Assistant G-3 Plans, were given appropriate excerpts from the corps OPORD which assigned the mission. Subjects then queried their computers to obtain such data as they felt they needed to complete three planning phases. Subjects completed the sequential phases by filling out response sheets. Response Sheet I required the subjects to record his allocation of combat power to the echelons of defense, Sheet II requested him to record his task organization, and Sheet III required him to write out missions for subordinate brigades. There were 20 subjects, 19 lieutenant colonels and one colonel. Data collected were the same as in the first experiment. The CGSC school solution was used as the criterion measurement.

Table 23 shows the correlations between nine predictor variables and the measure of how well each subject defined the solution. Data from the first experiment are also included. Clearly the schoolhouse concept that planning performance can be predicted from either data seeking behaviors or biographic variables suffers severely when the data from this experiment are analyzed. However, the research does demonstrate that a SIMTOS like research facility helps make research possible and that the schoolhouse solution, which is, of course, based on expert judgement, provides a suitable method for determining planning effectiveness.

Table 23

Product-Moment Correlation Between Nine Predictors and the Criterion

Predictor	First Experiment N = 20	Second Experiment N = 20
Recency of graduation from CGSC (EL)	.20	.60*
Experience in Mechanized Infantry (EI)	.43**	.05
Experience in exercises in Germany (EX)	.46**	-.19
Class Standing at CGSC (ALC)	.47**	.29
Expressive (CGSC Standard) (ALW)	.42**	-.02
Sequence Score (PSEQ)	.53**	-.07
Runs Score (PDRR)	.49**	-.11
Slope Score (PRS)	.44**	.35
Terminal Pause Score (PTP)	.22	.23

* Significant beyond .01 level.

** Significant beyond .05 level.

EL through ELW one-tailed, PSEQ through PTP two-tailed test.

Beyond these few studies no division level research is available. SIMTOS is gone now and apparently was not used for any additional command and control research relevant to the area of this review. The concept represented by SIMTOS has, advanced of course, and more elegant but equally dedicated facilities are becoming available. It may occur to the reader that the division ARTEP which was published in 1978 might have provided an opportunity for data collection at CPXs conducted to exercise division command and control operations. This does not appear to have occurred. Indeed, the most interesting thing about the division ARTEP may well be its course since its inception in 1978.

The reader might recall that the ARTEP process was developed to replace the earlier Army Training Test System (ATTS) which had been created by extending to groups the process used to create tests for individuals. The result had been bulky, rather detailed sets of test requirements. The original Division Command Group and Staff ARTEP, including appendixes, contained 57 pages. Today there are two division ARTEPS, a light division and a heavy division. The light division ARTEP, final draft, contains 822 pages of which 699 describe tasks and related material. The heavy division ARTEP, initial coordinating draft, prepared after the light division draft, contains 333 pages and six short appendices. Apparently evaluating performance at the division command post level is not so easily done that a clear approach exists even for so broadly applied a methodology as the ARTEP.

OTHER COMMAND AND CONTROL RESEARCH

Command and control research results of interest on the context of this review, but not already discussed, are entirely products of the application of a performance measurement methodology known as HEAT. HEAT, which is an acronym for Headquarters Effectiveness Assessment Tool, was developed for the Defense Communications Agency (DCA) during the period 1982 to 1984. DCA and the Defense Nuclear Agency (DNA) had been considering the problem of how to size, distribute, and equip theater level command posts in order to make them both survivable and effective. It was apparent to DCA and DNA personnel that in order to "size" a headquarters it was necessary to define the missions that the headquarters must accomplish and then to develop a method to measure the effectiveness of a command post staff performing the tasks needed to accomplish those missions. Clearly, unless a command and control performance measurement methodology of some sort was available the impact of staffing, configuration, automation, and other parameters that effect command staff performance could not be evaluated.

HEADQUARTERS EFFECTIVENESS ASSESSMENT TOOL (HEAT)

Since no methodology was available that even purported to measure command and control performance at the theater command post level funding was provided, to Defense Systems, Incorporated of McLean, Virginia, to develop a suitable methodology. Ultimately, DSI provided a series of reports which describe the HEAT methodology and the material that DSI's anonymous researchers¹ felt was relevant to understand and accept their process. The first report (DSI, 1982a) discussed headquarters measurement problems in general, the functions of a theater headquarters, and HEAT as a potential technique for measuring command post effectiveness. The second report discussed current and historic data on theater headquarters and problems associated with the need to measure effectiveness in such an environment. Two volumes of appendices, one secret and one unclassified, provided a considerable body of information on present and past theater command posts, and related topics. The third report (DSI, 1982c) was somewhat in the nature of a proposal and described certain issues which could be researched using HEAT.

The HEAT methodology is described in the projects fourth report (DSI, 1984), a HEAT User's Manual, which describes the methodology, its rationale and the process of applying it and evaluating the results. As described by the researchers HEAT is based on the concept that activities of command posts, or headquarters, are best conceptualized as an adaptive control process. In their model the researchers postulate six steps that constitute the phases of an adaptive control cycle. The six steps - monitor, understand, develop alternative actions, predict consequences, decide, and direct - constitute their sequence of each adaptive control cycle. The model, which is shown in Figure 12,

¹DSI publishes reports without individual authorship attributions.

defines three situations which can come into being as time and events go by and an OPLAN is compared to the perceived reality of the unfolding situation. These three situations are defined by the HEAT developers as being consequences of the divergence between the situation predicted by the plan and the situation that really develops. This difference, called "incongruence" in the HEAT terminology, may be minor, moderate or major.

A minor incongruence is one which can be compensated for with minimal adjustments considered to be within the scope of the original plan. A moderate incongruence represents a situation outside the scope of the existing plan but of such a nature that the possibility of its occurring had been considered. Thus, the developing situation had been considered as an alternative and predictions of various actions that could be taken if it occurred had been prepared. In cases of this type much of the work required to complete an adaptive cycle has already done and, only a portion of command and control cycle needs to be redone. A major incongruence is one that arises without suitable prior consideration that it could occur. Since data to support the understanding that the occurring events might happen either had not existed, or had been misinterpreted, the subsequent processes of considering the situation as an alternative and predicting the nature and impact of available responses had not occurred. In these cases there are no partial solutions to the command and control problem and the entire monitor through direct cycle must be traversed.

Three types of measures are defined in the HEAT methodology: effectiveness, process, and diagnostic. The HEAT developers faced an effectiveness measurement problem quite similar to the one faced by Olmstead and his coworkers a decade earlier when they did their first study on battalion command groups. Possibly because there were no battle outcome scores available to tempt them Olmstead and his coworkers, without specifically addressing the issue, opted for a decision quality based effectiveness measure. The HEAT developers were clearly in the same position. Since there were no theater combat models to provide easily available - but deceptive - battle outcomes to serve as effectiveness scores the HEAT developers had to specifically address the issue of command group effectiveness. Possibly because at least one of the senior investigators had done work on "unit" effectiveness as opposed to "headquarters" effectiveness (see Hayes, 1977), the HEAT team concluded that since the products of a headquarters are its plans a command and control group's effectiveness is best measured by the effectiveness of its plans. Plan based measures were, therefore, defined. The first measure is based on the percentage of intended period the plan remains in effect. To defend this measure HEAT developers articulated the reasonable concept that when a plan ceases to be effective in controlling the events it was created to control a new plan will be seen as necessary and then prepared. Hence, a command and control process is good if its plans do not fail and bad if its plans require early replacement.

The second effectiveness measure is based on how plans fail. The measure considers that if a plan must be changed it is reasonable to conclude that being able to replan without having to go through the entire planning cycle is better than having to go through the entire cycle. HEAT assumes that command groups that are able to do a greater quantity of planning tasks per unit time

are able to consider more understandings, develop a broader range of alternatives, and evaluate more predictions. It follows therefore that these better groups are more likely to have partially considered the evolving situation and, hence, need only to partially replan. Such groups also are more likely to have selected a better original plan than less able groups who must select from fewer and less well analyzed potential courses of action.

Process measures, there are four, are based on time. Three are cycle dependent measures: average time to complete a full planning cycle, average time to replan when only a partial cycle is required, and average time to make an adjustment to a still-in-force plan, (the three cycle types shown in Figure 12). The fourth measure is the percentage of orders or directives which arrive at subordinate headquarters early enough to provide adequate planning time.

The HEAT description contained in the HEAT Users' Manual, describes 135 diagnostic measures. The manual also notes, in describing how to plan and conduct a HEAT application, that not all diagnostic measures are appropriate to a specific exercise and, consequently, part of the planning process needs to be the selection of those that make the greatest contributions to the goals of the exercise or research event.

HEAT was developed for theater level command and control analysis activities. Indeed in one of their early reports the HEAT developers note that their methodology might be applicable as far down as corps headquarters. Elsewhere, in an unclassified Appendix to one of their reports (DSI, 1982b) the developers provide a comparison of selected headquarters. The data shows that theater command posts vary considerable in size. The variation is from several hundred (e.g., CINCPAC and early 9th AF organization) to many thousand (e.g., SHAEF, 1945). Since the developmental and most subsequent application of HEAT were at naval command posts, which were smaller than brigade level Army command posts, it seems quite likely that the concept can be modified and applied to a broader range of command post sizes than originally considered.

HEAT APPLICATIONS

Since its initial development in the 1982 to 1984 time frame HEAT has been applied in four general areas: to studies of naval command posts, in Naval Postgraduate School command post training exercises, to selected military operations data, and, more recently to division command posts. The applications to naval command posts and the division command post applications are classified and distribution has been limited. Three Naval applications (DSI, 1985a, 1985c and 1986a) are classified secret and have not been examined for this review.

Naval Postgraduate School Applications

Two contractor reports have discussed research, conducted at the Naval Postgraduate School, which used HEAT measures as part of a command post evaluation process. The nature of the three experiments reported (Naval Command Post) and the limited use of the HEAT methodology cause the work to be of only

marginal interest in the context of this review. In the earlier of two reports the researchers discussed two experiments one of which dealt with "connectivity" and the other with "organization." In the other report a third experiment dealing with the effect of command and "centrality" is discussed.

Experiment 1, Connectivity. In the first of these reports Strack (1985) reported on two experiments conducted at the Naval Postgraduate School WARLAB using officer-students as subjects. The first experiment, which dealt with connectivity, was conducted to determine how the communications links available between four physically separated command and control activities affected the decision process during simulated naval engagement situations.

The three tested communication networks, shown in Figure 13,¹ were used to connect four decision makers: a composite warfare commander (CWC), anti-air warfare commander (AAWC), anti-submarine warfare commander (ASWC) and anti-surface warfare commander (ASUWC). Each of these four decision makers (nodes) had specific functional responsibilities and some resources but also had the need to coordinate and share with other nodes. Communication among the nodes was restricted to an electronic mail network.

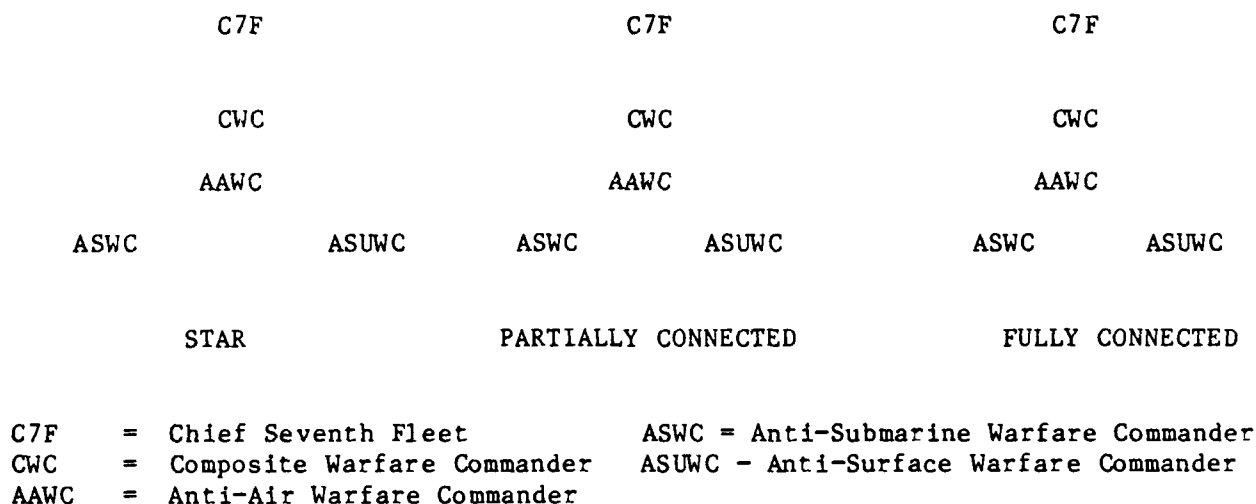


Figure 13. Headquarters Structures Tested at NPGS.

Three basic scenarios were used to provide each team of participants with three problems, differing in details but not in overall nature. All scenarios involved a battle group operating in the Sea of Japan and required it to strike a balance among its missions of sea control, power projection, and self-defense. In each case, sea control meant keeping sea lines of communication

¹The authors note that they developed their networks based on unreferenced Russian research. Their reference is probably Druzhinin and Kontorov, 1972.

open, despite interdiction attempts, during a conflict in Korea. Power projection required support of forces ashore in Korea and relief of OPFOR pressure, possibly including invasion on the Japanese island of Hokkaido.

The first part of each scenario represented a period of five days in which a carrier battle group, under command of the BLUE players, is en route to the scene of action off Korea. During this time, the BLUE team receives five successive daily "glimpses" of OPFOR force movement and activities, supplemented by situation reports from the fleet commander. BLUE can prepare plans for deployment, maneuver, and engagement upon arrival at the scene of action at any time during this period. The second part of each scenario began with the carrier force nearing its station. The subjects were given a background briefing covering the preceding five days, given a new daily glimpse, asked for a plan of action, and then given an opportunity to execute the plan.

The first five days of each scenario presented the BLUE headquarters with problems requiring the four command groups to formulate an understanding of the situation, develop alternative actions in response to that situation, assess the likely outcomes of these actions, and decide upon a plan for their forces. The activities of the sixth day, which could include simulated combat, provided an opportunity to judge each headquarters performance by the appropriateness of final decisions and, if applicable, by the outcome of combat.

The basic differences among the scenarios was in the underlying OPFOR intentions and in the military problem with which faced the battle group on Day Six. Scenario A confronted the battle group with the problem of self-defense against a deliberate attack, based on the OPFOR objective of isolating the Republic of Korea. Scenario B required the battle group to assist in a defense against an amphibious assault. Scenario C focused attention on defense of sea lines of communication with OPFOR intentions limited to direct support of North Korea.

OPFOR actions on Day 6 differed significantly from session to session to minimize "learning effects and possible unofficial communication among the teams between sessions." OPFOR behavior included one of the following: a deliberate full-scale attack; provocation, i.e., hostilities on a limited scale, which BLUE might attempt either to contain or to escalate; and intimidation, in which OPFOR conducted threatening maneuvers but did not attack unless attacked by BLUE forces. In all cases, OPFOR responded to an attack by BLUE with an all-out counterattack. Instructions to the battle groups were to defend itself against clear-cut attacks, but to undertake no offensive action until so directed by the Chief of Seventh Fleet. Counterattacks were permitted only against an attacking force, not against the OPFOR wherever found.

Each of the four warfare commander areas contained a graphics display, an electronic status board, an input terminal and an electronic mail terminal. Each group was aided by one or more "non-playing" assistants who manned the various input terminals. A total of six teams were tested and each team participated in three trials. Each trial involved a different combination of the three nets (minimum, intermediate, and fully connected), the three enemy sce-

narios, (interdict the sea lines of communication, intercept the battle group or invade Japan), and three levels of OPFOR threat to the battle group (intimidate, provoke, attack).

Data were collected on 13 (of the possible 135) HEAT measures. The measures used were:

1. Hypotheses about enemy objectives.
2. Number of hypothesis about enemy objectives.
3. Hypotheses about enemy assets assigned to objectives.
4. Number of Hypothesis about enemy assets assigned to objectives.
5. Hypotheses about enemy assets assigned to reserves.
6. Number of hypothesis about enemy assets assigned to reserves.
7. Overall ops plan for each contingent.
8. Component ops plan for each contingent future.
9. How many overall options for most likely future.
10. How many component options for most likely future.
11. Predictions about overall ops plans.
12. Predictions about component ops plan.
13. Average time to adjust ops plans.

Experiment 2, Organizational. In the later experiment reported by Strack in 1985 two organizational options were tested: functional and geographic. With a functional organization each command node directed a specific type of warfare. With a geographic organization each command node directed all warfare types within a specified geographic sector. The actual setting for the experiment was again the Naval Postgraduate School War Laboratory. The Warlab was configured into three battle group command nodes, each of which was equipped with a graphics terminal displaying a tactical situation map, an electronic status board with tactical status information, an input terminal to receive commands and control displays, and an electronic mail terminal. Two additional command nodes contained only a manual plotting board and the electronic mail capability. The five segregated areas represented CINCPACFLT (which was also control), COMSEVENTHFLT, overall Task Force Commander/Constellation Battle Group, the Kittyhawk Battle Group and the America Battle Group. The three carrier battle groups had the fully automated equipment suites. The specific setting for the experiments was a 3-carrier battle group operating in the Arabian Sea just outside the Strait of Hormuz. The battle group mission was to provide tanker convoy escort in the Persian Gulf and Hormuz Strait and to conduct self-defense as necessary. The situations used to test the subjects differed on four dimensions: organization, scenario, level of communication disturbance and OPFOR intentions.

The U.S. force mission was to escort tanker convoys in the Persian Gulf and Hormuz Strait and to conduct self-defense as required. The actual need for convoy protection and self-defense varied among a set of scenarios constructed for these experiments. There was range of culprits who could become involved as the OPFOR as the scenario unfolded: Iran, in retaliation for a Kharg Island attack; Iraq, to keep Iranian shipping at risk throughout transit; USSR to destabilize the area and embarrass the U.S.; Iraq and Iran in mutual escalation pursuing reasons cited above; Iraq and USSR, in a Soviet effort to enhance

their influence using a client state, and a scenario that had no developing OPFOR. The researchers also introduced two levels of communication (clear and disturbed i.e., with a receive capability deleted) and three levels of OPFOR intent (attack, intimidate, provoke).

Three groups were run and each group participated in six trials each of which was based on one of the six "culprits." Some attempt was made to distribute the 18 (of 72) variable combinations but at best the problem of too many combinations and not enough teams was difficult. Within the nodes the command team members rotated between the roles of commander, communicator, and "player (plotter at COMSEVENTHFLT)" during successive trials. Nineteen HEAT measures were collected during the trials. The HEAT based measures are shown in Table 24.

Table 24

HEAT Measures Used in Experiment Comparing Geographic and Functional Organizations

Effectiveness/Process	Measures
Overall Effectiveness	Plan life span Time to replan Exchange ratio, totals*
Monitor	Overdue reports Overdue reports not queried
Understand	Identify opponent Opponent goals Assets assigned to goals Opponent schedule Guidance understood
Options	Number of options Options for each contingency
Prediction	Expected results for each option Option vs contingency assessment
Decision	Response threat Shrink or deny OPFOR options Preserve or expand own options
Direct/Inform	Directive reflect decisions Queries Shared perception

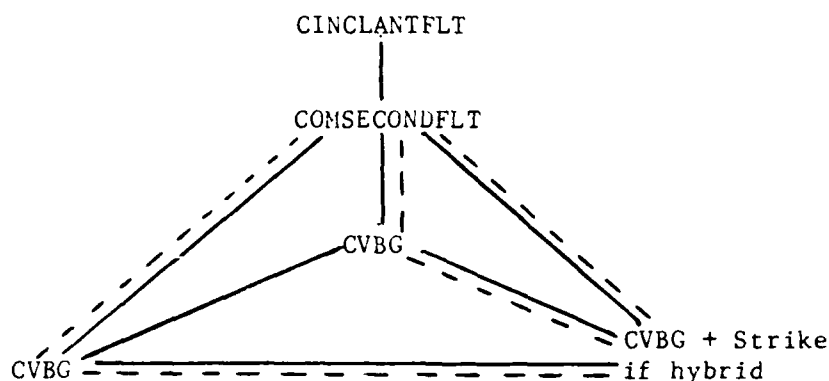
* This was a battle outcome measure not a HEAT measure.

Results

The author discusses the results of the two experiments at considerable length - 25 pages - but the section does not really address results in the manner considered necessary to the trained behavioral researcher. The most compelling characteristic of the section is that the HEAT based measures which were collected appear not to have been analyzed. Indeed most of the 13 HEAT measures taken in the "connectivity" experiment are not a part of any analysis and none of those taken in the "organization" experiment appear to have been considered. Further obfuscation occurs because the five measures labeled, "HEAT" in the analysis (actually four and "average HEAT") do not correspond to those listed in earlier sections as collected. The result of these omissions, is that no conclusions can be drawn concerning how well HEAT measures sufficed as measures of command and control performance. Despite these, and other, analysis problems the author provides an extended discussion of the relationship between his results and "command and control theory." In general these experiments appear to have been poorly conducted and ineptly analyzed. Taken in their entirety they suggest the need for a more rigorous approach and a more constrained extrapolation from what needs to be better data.

Experiment 3, Role and Specialization. A later experiment conducted in the Naval Postgraduate School, also used HEAT measures. As described by the researchers (see DSI, 1986b) this project was the third event in a series, each one designed to determine the impact of a basic command and control attribute --connectivity, centrality, and role -- on command and control effectiveness. This third event examined the impact of role and specialization. The experiment was conducted with the five cells described in the previous experiment which compared the performance of geographic versus functional organizations. This experiment compared the performance of higher and lower echelons (i.e., Fleet and CVBG nodes) in the conduct of planning and battle management tasks. It also compared battle force performance with the force configured in either a geographic organization or a hybrid combination of geographic and functional organization. In the hybrid organization, the CVBGs organized geographically with two exceptions: "Strike" planning and execution was centrally controlled by one of the CVBGs and another CVBG was responsible for AAW coordination. In the geographic organization each CVBG was responsible for all warfare areas within its geographic area, but the "Strike" function was coordinated at the Fleet node with execution being the responsibility of the individual CVBGs.

The distinction between geographical and hybrid organization is illustrated in Figure 14, which also portrays the experimental communications network. Within this network, the "Strike" node was collocated with a CVBG for the hybrid organization. For the geographic organization the "Strike" function node, which operated in a coordination role, was collocated with COMSECONDFLT commander. As in the previous experiments each CVBG command node included the following equipment: a graphics display, an electronic status board, an input terminal, and an electronic mail terminal.



_____ Warfare Area Communications
 - - - - - STRIKE Communications

Figure 14. Communications network used in role and specialization experiment.

The COMSECONDFLT node contained a communications terminal, graphics display on which any CVBG tactical situation display could be selected, and an input terminal. The CINCLANTFLT node contained only a communications terminal and relied solely on electronic mail for information. Each CVBG node was staffed by a three-person team: a communicator, a game player, and an action officer in overall charge of assessment and decisions. The COMSECONDFLT team consisted of a communicator, a plotter, and an action officer in overall charge. The node where the "Strike" was located was staffed with an extra player plus an extra electronic mail terminal. The CINCLANTFLT node was staffed by a single player, who was part of the control team. The individual players' roles were constant for the entire set of experiments and NPGS students who played the roles were assigned to them by staff members based on staff judgments of where the students would perform best.

For each experimental trial, the scenario was presented to the group, using a hard-copy SITREP to define the situation, followed by an interactive computerized wargame to present scenario development and allow them to respond to the problem as they perceived it. In all there were three independent variables which were recognized by the researchers prior to the experiment and a fourth, whose recognition after the experiment, indicates the behavioral research sophistication of the researchers. The recognized independent variables were: OPFOR intent (simple counterforce, 2-prong counter force, escorted counterforce, and attempted envelopment), organization (geographic and hybrid), and communications (clear and disturbed). There were two groups each of which completed nine trials. The researchers attempted to control "for the effects of learning, and to minimize impact from unauthorized crosstalk between groups" by their experimental design. Since the individual three hour sessions represented three prebattle days and followed by one battle day this was not a simple process. The geographic versus hybrid variable was simple enough but intent and communications were more of a problem. Enemy intent actually

permitted 16 different scenarios since each "intent" could begin a sequence and, either remain the same or change to one of the other intents for the day four battle. Thus, there were 16 possible combinations of OPFOR intent.

The communication variable was far more variable than the clear versus jamming dichotomy implies. Jamming which effects reception but not transmission could be applied to one or any combination of nodes in a wide variety of ways. The researchers decided to create communication disturbances for approximately ten percent of the available "circuit-minutes" during each of two required planning and action cycles on each of the four simulated action days. Thus, there was a potential jamming insertion for each of the four tested nodes and for a rather large set of combinations of the four nodes. The researchers opted to limit their jamming to one of the player nodes and not to jam the Control/CINCLANTFLT node which was actually a control function. The researchers, beyond their statement that they tried to control for learning and cross-talk, did not address how they felt their design coped with interaction problems associated with the independent variables. They did, however, provide a design which had each of the two groups see nine combinations.

Data collected by the researchers included seven HEAT based measures and data on over 100 parameters or incidents that were available from battle simulation outputs and/or from analysis of simulation recordings. These latter data types included such items as OPFOR commands issued, average messages to battle groups, and battle group radius of view.

Results. An interesting result of this study was that the researchers discovered the concept of individual differences and the potential impact of such differences on small group behavior. The authors note that, "The experiment design was based on the assumption that the two groups would be of like population and would react similarly to identical (or nearly identical) stimuli . . . (however) each group behaved in a distinctly different manner... As a result of these differences, overall performance scores tend to reflect an averaging of two different populations rather than a representation of a homogeneous population. Therefore, the analysis includes separate group performance scores along with the overall performance scores for comparison." Thus, another independent variable was added.

It would be unkind to dwell too long on the lack of statistical acumen demonstrated again in the handling of the available data. Most of the data, over 100 items, are manipulated relegated to appendices, and forgotten. In this experiment however, some of the HEAT scores are tabulated and used to draw certain conclusions. The tabulated HEAT scores, as provided by the authors, are shown in Table 25. It should be noted that in the HEAT methodology scores are converted to numerical values where 1.00 is the best score possible and .00 is the worst. These converted scores for four measures were then averaged to obtain the values shown in the table.

Table 25

Mean Values of Selected HEAT Scores

OVERALL - ECHELON		OVERALL - ORGANIZATION	
<u>FLEET NODE</u>	<u>CVBG</u>	<u>HYBRID</u>	<u>GEOGRAPHIC</u>
A = .39	B = .42	A = .37	B = .35
E = .35	F = .39	E = .31	F = .37
C = .43	D = .38	C = .32	D = .44
G = .30	H = .36	G = .30	H = .30

ECHELON - BY GROUP			
<u>FLEET NODE</u>		<u>CVBG NODE</u>	
<u>GROUP A</u> (Reactive)	<u>GROUP B</u> (Proactive)	<u>GROUP A</u> (Reactive)	<u>GROUP B</u> (Proactive)
A = .26	A = .42	B = .41	B = .46
E = .26	E = .63	F = .39	F = .38
C = .37	C = .48	D = .34	D = .44
G = .31	G = .29	H = .39	H = .32

ORGANIZATION - BY GROUP			
<u>HYBRID</u>		<u>GEOGRAPHIC</u>	
<u>GROUP A</u>	<u>GROUP B</u>	<u>GROUP A</u>	<u>GROUP B</u>
A = .38	A = .37	B = .30	B = .38
E = .31	E = *	F = .30	F = .40
C = .27	C = .40	D = .42	D = .45
G = .30	G = .28	H = .36	H = .27

LEGEND:

A,B = Planning phase, clear communications.

E,F = Planning phase, disturbed communications.

C,D = Battle management phase, clear communications.

G,H = Battle management phase, disturbed communications.

* = No data available.

The authors use the data in Table 25 to draw five conclusions. They state that the averaged HEAT scores indicate:

1. Group B (Proactive) performed better in the Fleet role than in the CVBG role.
2. Group A (Reactive) performed better in the CVBG role than in the Fleet role.
3. Group B generally performed better than Group A in a benign communications environment.
4. Group B was more affected by communications disturbances than Group A.
5. Group A was affected by the type of organizational structure while Group B was not.

The researchers did not report any attempt to subject their data to statistical analysis. Thus, it appears their conclusions may have been based solely on inspection of the data. A similar inspection by the reviewer provides no reason to believe that the data demonstrates any significant differences. An analysis, by the reviewer, of some of the data shown in Table 25 suggests that there is less to the conclusions than meets the eye. A Mann-Whitney U test was applied to the data which could be associated with some certainty to three of the five conclusions. Table 26 shows the comparisons and results. The researchers noted that they discovered that the groups were different after the data analysis began and that they had expected the groups to perform similarly in the test situations. A two tailed test was therefore performed and the .05 level accepted as significant. The data in Table CC clearly shows that there is no basis to claim that the differences noted in Conclusions 1, 2, and 5 actually exist. It appears that the researchers were overwhelmed by an over ambitious but inept experimental design, and never found a way to cope with their data analysis problems. This experiment, like its two predecessors, has limited value but does introduce another research topic which, properly exploited, might make a contribution to understanding the command and control process.

Related studies. Two other reports need to be mentioned at this point. The first is a Naval Postgraduate School Thesis by Owens and Brown (1984) which report on the same experiment already discussed by Strack (1985) and reviewed in section Experiment 1, Connectivity. The Owens and Brown thesis, completed as part of their course work, contains a coherent exposition of the background and conduct of the connectivity research which was jointly conducted by DSI and NPGS personnel. The authors treat the data obtained in the experiment in a more sophisticated statistical manner than did the DSI investigators.

Table 26

Mann-Whitney U Test Applied to Selected DSI 1986b Data

Conclusion	Relevant Data		Mann-Whitney U*
#1	42	46	6
	63	38	
	48	44	
	29	32	
#2	41	26	1
	39	26	
	34	37	
	39	31	
#5 ***	38	30	6
	31	30	
	27	42	
	37	36	

* A "0" is required for $p < .5$ two-tailed test with four pairs.

** The second part of conclusion 5 (that Group B was not affected by organizational structure) was really a trivial conclusion. Since there were only three data pairs and four is the minimum number required to obtain any level of significance.

In their results section the authors report that Mann-Whitney U tests applied to various experiments results found no case where there was a statistically significant (beyond the .05 level) difference between data sets from the three communication networks. They did conclude, however, that the results were consistent with the hypothesis that star structures are generally faster but that fully connected structures are more often correct when making creative decisions. The authors also determined that only three of the 13 HEAT measures used proved "serviceable". The usable measures were those derived from: (1) hypotheses about enemy objectives, (2) number of hypotheses about enemy objectives, and (3) average time to adjust operations plans.

The second report of interest (DSI, 1986c) is an effort that compares the results of the Warlab based experiments, discussed above, with results obtained from in-port training exercise results (DSI, 1985c, 1986a) and at exercise Bold Eagle (DSI, 1985a). The report attempts to make these comparisons in the context of a command and control theory articulated in earlier reports (DSI, 1983, 1985b). The most interesting material in this report is found in a table which lists the HEAT measures which have been used in exercises and experiments. It has already been noted that in the experiments a variety of HEAT measures were

taken but that many were not included in later analysis attempts. Tabulated data in this report indicates that 30 measures were "used" in exercises but only five were used in experiments.

Thus, the report is of interest because it is an attempt to compare real exercise results to experimental results and to describe performance measures and other data, in the context of the DSI command and control theory. Though interesting, the attempt is at best weak since it relies heavily on data gathered in experiments which provided data that did not support the demands put upon them by researchers who seem to have allowed their expectations to exceed their data. When these kinds of results are then used as a basis for experiment to experiment and experiment to exercise comparisons, the results though interesting, are not particularly robust.

Division Command Post Applications

The HEAT methodology has been applied to two division CPXs in order to collect data to determine the value contributed to the command and control process by the Maneuver Control System. One exercise was at the IID at Fort Riley, Kansas (DSI, 1986d) and the second was at a IAD CPX in Europe (DSI, 1987). The data collection plan for these exercises is described in an MCS Handbook (DSI, 1986e). The process of collecting data during a division CPX is not insignificant. At the IID CPX, two observer teams were required to provide 24 hours coverage. Each shift required a team of observers stationed as follows: two at control, five at DMAIN, one at DTAC, and one at each brigade headquarters. Data were collected to support estimates for 33 measures. The measures are shown in Table 27, and it will be noted that five were "effectiveness" measures, six were used to measure to what extent back up or alternative nodes were prepared to take over another nodes functions and 22 were diagnostic measures.

Essentially the same measures were considered in the later IAD CPX evaluation. The actual results of both the IID and IAD command and control evaluations and the MCS versus non-MCS comparison data are both classified and beyond the scope of this review. However, it is not a violation of either security or commitment to unit personnel to report that the HEAT based evaluations were successful. Success being judged by the fact that tabulated data were sensitive enough to imply differences between locations, times, and/or in situations where differences could logically be expected to exist. For example, during the CPX one brigade had troops in the field taking ARTEPS and a second brigade had only a command post with support personnel and a link to control. Typically experience shows, and the HEAT application measured, a performance difference with the no-troops headquarters providing a better reaction to the division command post than the brigade with troops in the field with real world duties.

Table 27

Description of Measures Defined for MCS Evaluations*

Effectiveness Measures

- Percentage of intended period plan is in force.
- Percentage of control cycles arising from minor, moderate, or major incongruence.
- Average time used to complete the control cycle given full planning, minor incongruence, adaptive behavior (surprise).
- Percentage of directives for which planning lead time provided to subordinates is inadequate.
- Percentage of adaptations not attempted due to insufficient lead time.

Monitoring Measures

- Percentage of units where headquarters data are outside desired window.
- Percentage of units where most recent data are outside desired time window.
- Percentage of over age data not queried by HQ.
- Percentage of units located with good, adequate, and inadequate accuracy.
- Percentage of plans changed because HQ perception did not match truth.
- Percentage of commanders information requirements for which there is no data.
- Percentage of information received but not accessed by staff because of inability to access; data not stored; data dumped.
- Percentage of units where reported data on units are outside the desired accuracy window.
- Percentage of reports with good, adequate, and inadequate accuracy.
- Percentage of units where reported data are outside the desired time window.

Understanding Measures

- Percentage of intended period hypotheses about enemy situation are: correct, not correct, or incorrect.
 - Percentage of time there is no hypothesis about enemy situation.
 - Percentage of attempts made to implement a HQ decision in which the match between the predicted and actual own force situation is wrong or left unmade.
 - Percentage of plans changed because situation was misforecast.
 - Percentage of intended period reported hypotheses about units are: correct, incorrect, or never reported.
-

Table 27 (Continued)

Description of Measures Defined for MCS Evaluations.*

Alternatives Developed Measures

- Percentage of all alternative futures hypothesized in "Understood" examined during the planning process and alternative courses of action planned for each contingency.
- How many people participated in the development of options for dealing with the "future"?
- How many options were examined for dealing with the future considered to be most likely to occur?

Prediction Measures

- Percentage of intended period predictions are: correct in the contingency set, or incorrect.

Direction Measures

- Percentage of assignments in directives which contradict the assignment of the commander's decision(s).
- Average time taken to issue a directive after a decision has been made.
- Average number of queries received per directive.

Alternate/Backup Node Comparable Measures

- Alternate and backup nodes have current information, adequate capacity, and no conflicting assignments.

Monitoring Accuracy Comparison

- Percentage of hypotheses about enemy objectives that coincide, overlay significantly, or differ.
 - Percentage of attempts made to implement a HQ decision in which differences in perception about own force situation cause queries.
 - Percentage of nodes incorrectly identified or not identified.
 - Percentage of directives and/or guidance issued by alternate and/or backup nodes which are redundant or conflict with existing directions.
-

*There were also seven measures defined which were specific to MCS. These measures are not relevant to this review.

HEAT EVALUATION

It seems that the applications of HEAT derived performance measures in experimental settings have been less than satisfactory. In some studies the measures seem to simply drift away and, although taken, they do not appear in the analysis. In other reports merged and manipulated remnants of a more numerous set of measures remain but lead to insignificant results. It has already been noted that much of this problem is apparently the result of inadequate experimental design and lack of suitable statistical acumen on the part of the researchers. In fairness, however, it must be noted that the kind of research being done was of such a nature that even with better research skills problems would probably have arisen.

Attempting to determine the relationship of individual measures to parameters that are beyond some direct command post effectiveness measure has been difficult even for fully trained, sophisticated researchers such as Barber, Kaplan, Olmstead, and their associates. Recall that in attempting to determine how ARTEP tasks related to battle outcome these experienced researchers found that individual tasks correlated only weakly to battle outcomes and they ultimately found it necessary to apply advanced statistical techniques such as factor analysis to their data. Even then conclusions often were that the data was not sufficient to demonstrate direct ARTEP task to force effectiveness measure relationships.

Probably the best way to define the potential of a HEAT based set of measures as research and operational command and control performance measures is by comment on a formal evaluation of HEAT, conducted by behavioral scientists at the Naval Personnel Research and Development Center. The letter report, which is included in its entirety in Appendix N, addresses both the strengths and weaknesses of the approach. As expressed by NPRDC evaluators, "The advantages of the HEAT methodology are: (1) its systematic model-based nature, (2) its comprehensive assessment of planning functions, (3) the quantitative measures it provides, (4) its use of multiple incidents on which to base performance scores, and (5) the lack of alternative assessment methodologies." The disadvantages of the HEAT methodology are: "(1) its focus on process at the exclusion of outcomes, (2) the lack of assessment of platform performance and execution of tactical decisions, (3) the subjectivity of measures and performance standards, (4) the quantity of data collection and analysis required, and (5) its unknown psychometric qualities."

Since there is some reason to believe that disadvantages, 1 and 2 are neither disadvantages nor even undesirable features for command post performance measure we may set them aside at least for now. It can also be argued that this methodology is less subjective than are ARTEP evaluations and that no methodology currently applied to command and control has known psychometric qualities. Thus, it appears that the existing methodology has promise and with careful development might become the measurement technique of choice for certain command and control evaluations.

SUMMARY

Researchers in the command and control area have explicitly or implicitly addressed an assortment of issues in the area. Certainly the problem of what is the best way to measure command group effectiveness has been addressed. Attempts to deal with how battle outcomes may be related to command group performance have developed from this issue. Effectiveness measures based on the quality of some command post product have also been developed with the result that meaningful approaches based on decision quality and on plan quality have been reported. ARTEPs, and issues related to them have been responsible for a considerable portion of the available literature.

ARTEP based research has addressed such factors as: Do simulators support ARTEPs? How does task performance relate to battle events? How well do raters perform their functions? Researchers have also addressed, by their conduct of ARTEP based research projects, the issue of how well can command and control research be conducted in situations where the primary purpose of the observed event is such that it can, and often does, seriously degrade the accuracy of the researchers data.

COMMAND POST EFFECTIVENESS. WHAT TO MEASURE?

A problem which is constantly faced by researchers in the command and control arena is what to use as criteria to specify the effectiveness of a command and control process. One of two, quite different, approaches is usually adopted when the problem of assigning some sort of an effectiveness score to a CP training or research event is addressed. In research projects previously discussed both approaches were used. The two approaches seem to relate to how the researcher, developer or trainer perceives the unit whose effectiveness he wishes to measure.

Often the researchers have accepted the concept that command and control effectiveness is synonymous with unit effectiveness. Thus, if the project deals with a battalion command post the researchers or trainers opt for a measure or measures based on battle outcome. Other persons facing the same problem take the view that a command and control organization is not synonymous with the organization it controls. Hence, they try to develop measures of command and control effectiveness that apply specifically to the command post as an entity.

One can somewhat facetiously describe these two positions as the "It's not how you play the game, it's if you win or lose that counts" versus the "It's not if you win or lose, it's how you play the game that counts" schools of thought. It is a serious matter, however, made more serious by the fact that persons often accept one model without any apparent awareness that the alternate model exists. The tendency has been to accept the former model without explicitly considering the latter model or the data which strongly suggest that

better research, better training, and, perhaps, even better doctrine would result if the latter model were used. The former view may have predominated because win-lose scores of some sort are the natural, and proper, outcomes of battle simulators that include models of force-on-force combat. This in turn causes the developers and users of such simulators to be very ardent supporters of those measures and to present them as the only necessary measures of command and control effectiveness.

Two research projects by Olmstead and his associates, discussed in the sections entitled, "In the Beginning . . ." and "Olmstead Revisited" are probably the best examples of the two positions. In the earlier study (Olmstead, Christensen and Lackey, 1973) the researchers developed an excellent research model and used a command and control group specific measure for defining the organizational effectiveness of the staff groups under test. In the latter research (Olmstead, Elder and Forsyth, 1978) the researchers used much the same research model but used a battle outcome effectiveness measure.

A "Way You Play the Game" Example

The Olmstead, Christensen and Lackey (1973) research created an effectiveness measure that was limited to factors that were almost entirely influenced by activity that was under the direct control of the personnel of the organization being evaluated - the battalion commander and his staff. In their 1973 work Olmstead and his fellow researchers obtained their effectiveness measure by a rigorous process of creating decision compelling events and measuring the quality of the resulting decision.

This approach assumes that the effectiveness of the battalion command and control process is best measured by examining the quality of the decisions made by the commander and his supporting staff. The model that supports the research also defines the battalion staff role in terms that make it clear that the command post must convey the decision to the appropriate higher, adjacent and lower units; must arrange to get the best available data on which to base their decision; must monitor the actions taken by subordinate units; and must perform the internal command staff actions necessary to maintain the command post's ability to continue to operate effectively in the battle situation.

An "If You Win or Lose" Example

A particularly interesting example of the type of research effort that uses battle outcome as a measure of command and control effectiveness is later work of Olmstead, Elder and Forsyth (1978). The work was directed by the same senior researcher who conducted the previously discussed example of a "how you play the game" approach. Since the research was based on much the same conceptual framework, the work serves extraordinarily well to demonstrate the differences in approach and results that characterize battle outcome oriented command and control performance effectiveness measurement methodologies.

In this latter study there were two major differences in the methodology used by the researchers. One was that combat outcome measures were used in place of the "best decision" measure used in the previously discussed re-

search. The other was that Organizational Competence was measured by ARTEP observer based subjective evaluations rather than by extensive analyses of the total staff communication process.

The effectiveness measures, designated as "Combat Effectiveness" measures by the authors, were based on the method suggested by Tiede and Leake (1971) in which two dimensions were identified that define the mission space of a unit. The parameters were Area, the area or geographical objectives controlled after completing the battle, and Resources, the quantity of resources (personnel, weapons, equipment) expended, or remaining, at the end of the engagement. Appropriate evaluations of the combinations of these two dimensions made it possible for the researchers to order participating units in terms of battle outcome.

There are two interesting aspects of the Olmstead, Elder, and Forsyth (1978) research. One is that although they did not address the issue specifically the researchers made a clear distinction between the command and control effectiveness concepts they used in their earlier research and the unit effectiveness measures used in the research now under discussion. In their earlier work the researchers considered their effectiveness measure to be specifically addressed to command group performance and, indeed, titled it Organizational Effectiveness. In this later project the researchers clearly felt that their measure was not command group specific and so indicated by their use of the title "Combat Effectiveness". Other researchers who have used battle outcome results (primarily, it seems, because battle simulation models provide a sizable number of such measures) have not been so careful - or so astute. Hence, as the reviewer noted earlier some of the researchers seem to have adopted the "win-or-lose" concept without considering, or even being aware of, the other concept.

The second interesting item concerning the Olmstead et al (1978) work was the relationship shown to exist between how competently the command staff performed and how "effectively" it performed. In the 1978 study the overall Organization Process scores correlated .63 with the area controlled/resources expended measure and .67 with the average of the mean of five ARTEP controller combat effectiveness ratings. The results suggest that the subjectively obtained Organizational Process scores were less adequate than the more rigorously derived measures used to quantify organizational competence in the Olmstead et al 1973 research. The results also suggest that factors beyond the direct control of the battalion command staff affect the strength of the relationship between command staff performance and battle outcome. It appears that the weaker organizational process quantification methodology and the insertion of intervening variables, implicit in the move from decision quality to battle outcome, resulted in the creation of a situation where the variance accounted for by the relationship between the command and control performance and the effectiveness measure used was much reduced. In the earlier Olmstead, Christensen, and Lackey (1973) work the relationship of Organizational Competence to Decision Quality was .93 which accounted for 86.5% (.93 x .93) of the variance. In the 1978 study Organizational Process accounted only 39.7% of battle outcome scores (.63 x .63).

RESEARCH ON BATTLE OUTCOMES AS MEASURES

Mention has already been made of the work of Tiede and Leak (1971) which supported the mission accomplishment/resources remaining measurement scheme which was used later by Olmstead, Elder, and Forsyth (1978). After 1971, research that addressed battle outcomes as effectiveness measures was often motivated by the availability of battle outcome data from evolving battalion level simulations. Thus, Barber, McGrew, Stewart, and Andrews (1979) when they conducted research to determine the training effectiveness and research vehicle suitability of CAMMS were also able to consider how battle outcome measures related to one another and to staff performance scores based on ARTEP evaluations.

The Barber, et al report contains a large number of correlations with all the problems that are associated with such statements as "The resulting correlation matrix with 376 entries contained 47 correlations significant at $p < .05$." Buried within their discussion of their results, however, is a discourse that suggests that they were addressing the question of "How are battlefield results most properly related to command staff performance"? The researchers analysis and discussion relevant to the what-to-measure question proceeded generally in the following manner. First the relationships among the various objective and subjective measures of battle outcome were analyzed and discussed. Then the relationships of various battle outcome measures to both subjective and objective measures of command group performance were explored. Finally the relationship of the objective battle outcome measures to controller assessment of ARTEP subtask performance were considered. The correlations among five subjective battlefield outcome measures and three objective battlefield measures suggested several conclusions concerning the measures and the command and control outputs that lead to battle events. Noteworthy was the fact that the computer generated personnel and equipment losses correlate very highly with each other, suggesting that the algorithms presume attrition on these two dimensions to be closely linked.

The Barber et al research also illustrates another effectiveness criteria problem. In their research the battle losses of men and material and "area lost" correlate highly, but negatively, with "enemy thrust," which is an estimate of the magnitude of error in identifying where the major penetration will occur. This, according to the researchers, suggests that friendly personnel, equipment, and area losses will decrease as the error in estimating where the enemy will be increases. This situation was not considered to be unreasonable for a covering force operation. The authors felt it meant that an error in locating the enemy thrust would reduce the frequency and intensity of fire exchanges and, thereby, decrease personnel, equipment and area losses. The substantial negative correlation between area lost and number of decisive engagements also seemed appropriate in that more frequent decisive engagements might in the short term reduce the area given up even though it results in less adequate mission accomplishment. According to the researchers the absence of a correlation between number of decisive engagements and mission accomplishment tends to confirm this reasoning. The researchers note that data like these demonstrate a recurring problem associated with using battle outcome as a command and control effectiveness by raising the question, "How does a researcher

explain good performance leading to bad results or bad performance leading to good results?"

The Barber et al (1979) study is also of interest because of the manner the researchers approached their evaluation of how battle outcome results relate to a command group's "organizational competence" as measured by ARTEP task performance. Barber and his associates did not appear to hold the belief that there was some battle outcome measure, or set of measures, that could serve as a command and control effectiveness measure. In discussing their research the authors stated that, "the relationship of command group performance to battle-field outcomes is complex," and that "the quality of performance on no single measure yet identified can be adequately interpreted in isolation of other measures, or the condition of performance." They also noted that although performance on some ARTEP subtasks does appear to influence battlefield outcomes additional effort would be required to obtain reliable estimates of the degree to which performance on ARTEP subtasks relate to outcome measures.

This research and other similar studies tend to support the concept that measures of the competence of a command and control organization can best be validated by demonstrating that they correlate with command and control effectiveness measures that are command post specific: such items as making the right decision, preparing orders that last, doing the doctrinally correct thing in task organizing or assigning missions. The broader question of how these command post specific things relate to battle outcome is much more complex and incorporates such questions as: "Is doctrine in this area correct?" "How does company training or moral affect the course of a battle?" and "What can the enemy do to counter my best laid plans?"

Other work has also shown that battle outcome measures are troublesome. Thomas (1983) concluded that his results demonstrated that some degree of validity could be attributed to simulation outcomes as measures of command group performance in CATTs exercises, that those calculations were sensitive to changes in mission and unit type, and probably to differences in pre-battle combat ratios. Thomas felt his results were particularly promising due to high correlations between simulation outcomes and controller ratings of command group performance. Thomas concluded that further research on the development of indices of simulated battle performance should include controlled or systematic manipulations of the above variables along with controller influence and terrain factors and that new indices should include additional components of mission accomplishment to supplement the simulation outcomes used in his research.

But Thomas also pointed to problems in his research: (a) correlations that might be inflated because of the way the data were obtained, and (b) the controllers', who were also players may have affected battle outcome. It is interesting that although Thomas' research was intended to determine the value of battle simulation outcomes as measures, he ultimately concluded only that he could develop no mutually agreed upon measure and that additional research should be directed at identifying such a measure.

In a later attempt to develop a composite measure of mission accomplishment, Thomas and Cocklin (1983) tried again to determine the feasibility of formulating a composite measure of mission accomplishment. The researchers added measures of various mission objectives which could be used in place of expert judges to assess battlefield performance in computer-driven battle simulations, such as CATTS/ARTBASS. In the end, the authors concluded that while it appeared possible to capture the decision policies of military experts in the form of a multiple regression model, a mission accomplishment score still needed to be more complex than a simple loss exchange ratio (and even more complex than their model).

PRODUCT QUALITY BASED EFFECTIVENESS MEASURES

Considerable attention has been given to the development of methods which permit the researcher to evaluate the quality of some command and control product and to obtain thereby a measure of the effectiveness of the process. Five such developments have been reported either as specific project goals or as incidental but necessary parts of a research development process. Stated in chronological order of development the measures are decision quality, school house solution comparison, OPORD completeness, OPORD briefing adequacy, and plan life.

Decision Quality

Olmstead and his fellow researchers provided the decision quality approach in their 1973 report and since it has been discussed in detail in earlier sections it requires no further discussion. Research reported the same year, 1973, by Krumm, Robins, and Ryan has not been discussed in such detail. In their project Krumm and his coworkers developed an experimental design, and a scoring technique, that permitted them to present a division level command and control problem and then score a subject's solution according to how closely the solution matched the Command and General Staff College solution. In a sense this was an advance on the concurrently evolving Olmstead work since the "schoolhouse solution" model provided a way to quantify the many decisions implicit in a division OPORD.

Neither of these approaches is easily applied. Olmstead's, as was noted earlier, requires considerable effort to develop probes and probe scores. The Krumm, Robins, and Ryan approach depends at best on having a scenario for which a schoolhouse solution is available, willing experts to provide scales, and weights for deviations that might occur on various parameters. At worst, it also requires the development of new scenarios along with their best solutions. Both approaches are likely to be difficult when attempts are made to apply them in dynamic division level simulations. Olmstead's probes (now missing, incidently) worked well in military incident list (MIL) based battalion level battle simulations. Krumm's approach worked well in decision making studies where periodic "command direction" moved deviant decisions back onto the proper course so that ultimately all decisions could be considered within the boundaries of the scales. Despite the difficulties that may develop, both methods are promising bases for command and control measurement processes that can consider the quality of the command and control product.

Product Evaluation

Two research groups have developed measures that provide a structured evaluation process for specific command post products. In both cases the measures involve the battalion OPORD, and since the battalion OPORD is generally briefed to subordinate commanders, both involved processes that aided evaluations performed as briefings occurred. The Archer, Fineberg, and Carter (1984) approach provided 15 areas which observers needed to rate for clarity, information content and quality of the proposed military solution. Ratings were done on a five point scale. The other approach, by Metlay and his associates in 1983, was similar but apparently less difficult. In the Metlay et al approach 30 items were defined, 22 were yes/no items, six were high/low items, and the other two only required the observer to provide a count and a time that could later be compared to standards. The relative success of approaches such as these suggest that in cases where observers must make on the spot judgments aids can be developed to assist them.

The final product evaluation item is the concept described in the HEAT methodology. In this methodology plan adequacy is inferred from the planning organization's implicit judgment of a plan's effectiveness and by the nature of the process that occurs when a new plan is required. In the HEAT scheme plans are good if they work and, therefore, are not seen as needing to be re-done by the command group. Also plans are better if they fail in a situation where only a partial planning cycle is needed to develop a replacement. This measurement approach appears very promising for dynamic simulation situations where command post activity is dictated by the commander's views of the developing battle situation. In these types of situations the concept that someone might stick with a bad plan in order to get a good score (while perhaps unnecessarily losing his unit) is not particularly viable. This measurement concept also helps correct for the other type of problem where a good plan is changed unnecessarily. Obviously an unnecessary change represents poor command group performance and quite properly reduces the erring command group's score in cases where a different system might only record that the old plan was a good one and that the replacement was also a good plan.

ARTEP RELATED

A major portion of battalion command and control research was motivated by the needs of persons developing simulators and the ARTEP system that was concurrently evolving. Researchers, always alert for an opportunity to collect data, were also able to conduct meaningful command and control research events while investigating the ARTEP/simulator arena.

Simulators and ARTEPs

A variety of studies have addressed the potential of the evolving battalion battle simulations to provide a vehicle for conducting ARTEPs. Kaplan and Barber (1979a) assessed comparative effectiveness and cost in a study that compared CAMMS based versus conventional CPXs. In the same period Barber, et al (1979) and Olmstead, et al (1978c) reported on other CAMMS based battalion and

brigade research involving ARTEPs. Both efforts demonstrated that command staff ARTEPs could be conducted on CAMMS but that not all tasks and subtasks could be evaluated. The problem of non-scorable subtasks was often the result of the tasks themselves and, less often, of not being playable on CAMMS.

Later reports such as Barber and Kaplan (1979) and Kaplan and Barber (1979b) provide data that supports similar conclusions regarding a competing battalion level simulation, CATTS. Still later, ARI command and control researchers used data, obtained on CATTS, to demonstrate that the evolving battalion combat simulation known as ARTBASS would be a cost effective and satisfactory training system. (Thomas, Barber, and Kaplan, 1984; Thomas, Kaplan, and Barber, 1984).

As these projects were conducted and reported it became clear that battle simulations were valuable training situations and that, although there were some problems, ARTEPs could be conducted using simulations rather than traditional CPXs. Other analyses conducted on data from these and other projects of the same type did provide some disquieting results. Most notable were continuing indications that better ARTEP task performance did not always assure better battle outcome results. A second problem was that the ARTEP raters did not appear to be good scorers of individual ARTEP subtasks. Raters, who were usually also controllers, noted that they were too busy to give the rating process the attention it required. Attempts to provide scales, better definition of ratings and other aids were never fully successful. Generally, individual subtask ratings were too much correlated with one another and clearly contaminated by the fact that the controller scores tended to be influenced by the state of the battle which was being fought.

In the end, conducting command and control research in ARTEPs performed in battle simulations was shown to be possible even though these situations did not provide a good research environment. It also became clear as time passed that sound statistical analyses were needed to capture the meaning of the resulting data, and that attempts to relate ARTEP tasks to battle outcomes were tenuous at best. Together these conclusions supported the concept, accepted earlier in the development of SIMTOS and articulated later by a series of ARI funded research efforts, that good research required a dedicated facility where researchers could control the command group's activities and the data collection process.

Ancillary Research

Despite the problems associated with piggy backed experiments, some very interesting research has been conducted in ARTEP/simulation situations where researchers were able to develop hypothesis which could be addressed by data naturally flowing from the events, or create relationships which enabled them to superimpose a research project on the training event. The Olmstead et al. (1978a) brigade level material is a good example. Olmstead was able to modify and expand his earlier organizational process/organization effectiveness model and develop evidence to support earlier findings that increased competence of the command group in performing its tasks leads to improved effectiveness in selecting better courses of action. Kaplan, in his 1980 report on information

flow, demonstrated that with good cooperation a research event could be run in conjunction with command and control evaluations intended primarily for some other purpose.

Later research by personnel at the ARI Field Unit at Fort Leavenworth also showed that data from only partially controlled events could be used in attempts to test command and control related hypothesis. It is probably safe to conclude that the numerous reports that tried to do science with data taken from beyond the behavioral science laboratory have successfully demonstrated concepts already held by many researchers. One verified concept is that great care needs to be taken to secure the willing cooperation of trainers, or other non-researchers, who have goals that may be impacted by a superimposed research goal. A second conclusion evolves out of the difficult nature of command and control research and the fact that additional difficulties are imposed when attempts are made to inject such research into some other event. Certainly the body of research by Barber, Kaplan, Olmstead, and their various associates also demonstrated, as time went on, that they felt the need to apply ever more sophisticated statistical techniques (e.g., factor analysis rather than mere correlations). Other research has convincingly demonstrated that inept experimental design and naive data analyses can lead either to failed experiments or improper conclusions.

Overall it seems that data from training or field events have the potential for making major contributions to the knowledge of command and control. But, at the same time it appears that the availability of such data places great demands on researchers to assure that sound scientific methods are applied to such material in order to prevent the data from becoming a source of misinformation. It also is clear that there is a place for dedicated facilities and a greater voice for researchers in planning training events or field tests since the cost of command and control research at higher echelons, e.g., division CP level, makes it imperative that the best use be made of every research opportunity.

References

- Altman, I. (1966). The small group field: Implications for research on behavior in organizations, in R.V. Bowers (Ed.), Studies in Behavior on Organizations: A Research Symposium, University of Georgia Press, Athens, Georgia.
- Archer, M. A., Fineberg, M. L., & Carter, C. F. (1984). Command group behavior, Objective 1, Final Report: Identification and analysis of battalion command group behavior. Science Applications International Corporation, McLean, VA.
- Army Training and Evaluation Program, Division Command Group Staff, No. 100-2. Headquarters, Department of the Army, Washington, D.C. 15 June 1978.
- Army Training and Evaluation Program for Maneuver Brigade Command Group and Staff, No. 100-1. Headquarters, Department of the Army, Washington, D.C. 22 May 1978.
- Army Training and Evaluation Program for Mechanized Infantry/Tank Task Force, No. 71-2. Headquarters, Department of the Army, Washington, D.C. 17 June 1977.
- Army Training and Evaluation Program, Mission Training Plan for Light Infantry Division, Command Group and Staff. No 71-101, MTP Final Draft. (Undated)
- Army Training and Evaluation Program, Mission Training Plan for Division. No 71-100 MTP Initial Coordinating Draft. (Undated)
- Barber, H. F., & Kaplan, I. T. (1979). Battalion command group performance in simulated combat (ARI Technical Paper 353). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A070 089)
- Barber, H. F., McGrew, J. F., Stewart, S. R., & Andrews, R. S. (1979). The computer-assisted map maneuver system: A preliminary examination of its training effectiveness and suitability for its use as a research vehicle (ARI Research Memorandum 79-9). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A077 981)
- Barber, H. F., & Solick, R. E. (1980). MILES training and evaluation test, USAEUR: Battalion command group training (ARI Research Report 1290). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A109 192)
- Bennis, W. G. (1966). Changing organizations: Essays on the development of evolution of human organizations. McGraw-Hill Book Company, Inc., New York.
- Bertalanffy, L. von (1956). General systems theory. General Systems Yearbook, Vol. 1, 3.

- Carter, C. F., Lockhart, D. C., & Patton, M. S. (May 1983). Command group behaviors: Their identification, quantification, and impact on collective output in automated and non-automated environments. Objective 1 Identification and analysis of battalion command group behaviors. Science Applications, Inc., LaJolla, California.
- Cooper, M. A., Shiflett, S., Korotkin, A. L., & Fleishman, E. A. (1984). Command and control teams: Techniques for assessing team performance. Bethesda, MD: Advanced Research Resources Organization.
- Daniel, C. (1976). Applications of statistics to industrial experimentation. New York: John Wiley & Sons.
- Defense Systems, Inc. (1982a) Theater headquarters effectiveness: Its measurement and relationship to size, structure, functions, and linkages; Vol I: Measures of effectiveness and the headquarters effectiveness assessment tool. McLean, Virginia.
- Defense Systems, Inc. (1982b). Theater headquarters effectiveness: Its measurement and relationship to size, structure, functions, and linkages; Vol II: Design considerations and guidelines for theater headquarters effectiveness; with classified (Secret) and unclassified appendices. McLean, Virginia.
- Defense Systems, Inc. (1982c). Theater headquarters effectiveness: Its measurement and relationship to size, structure, functions and linkages; Vol III: Approaches and methods for the validation of determinants of theater headquarters effectiveness: Proposed research. McLean, Virginia.
- Defense Systems, Inc. (1983). Headquarters effectiveness program summary, Task 002.
- Defense Systems, Inc. (July 1984). The headquarters effectiveness assessment tool: User's manual, McLean, Virginia.
- Defense Systems, Inc. (1985a). Assessment of headquarters performance in Bold Eagle 84 (U). McLean, Virginia.
- Defense Systems, Inc. (1985b). Elements of C² theory. McLean, Virginia.
- Defense Systems, Inc. (1985c). Second fleet battle force in-port training 2-85 HEAT analysis (U), Secret. McLean, Virginia.
- Defense Systems, Inc. (1986a). Second fleet battle force in-port training 1-86 HEAT analysis (U), Secret. McLean, Virginia.
- Defense Systems, Inc. (1986b). 1985 Command and control effectiveness experiments. McLean, Virginia.
- Defense Systems, Inc. (1986c). Correlation of experimental and exercise results. McLean, Virginia.

- Defense Systems, Inc. (1986d). Final performance assessment report First Infantry Division (Mechanized) Casus Belli VII (U), Confidential. McLean, Virginia.
- Defense Systems, Inc. (1986e). MCS HEAT Handbook. McLean, Virginia.
- Defense Systems, Inc. (1987). Assessment report: Vol 1. Initial determination of the value added by the MCS to Army command and control and Vol 2. Appendix D (U) Secret. McLean, Virginia.
- Donner, M. (1981). The role of information and instructions on the relationship between organizational communication and performance. Unpublished doctoral dissertation, Hofstra University, NY.
- Druzhinin, V. V., & Kontorov, D. S. (1972). Concept algorithm, decision. Washington, DC: U.S. Government Printing Office
- Evangelista, R. (1982). Organizational communication and decision making: A curvilinear model. Unpublished doctoral dissertation, Hofstra University, NY.
- Frank, E. (1981). Communication effectiveness: A multivariate analysis of performance, group structure and communication. Unpublished doctoral dissertation, Hofstra University, NY.
- Garlinger, D. K. & Fallesen, J. J. (1988). Review of command group training measurement methods (ARI Technical Report 798). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.
- Hayes, R. E. (1977). Measurement of combat effectiveness in Marine Corps infantry battalions. Technical Report Phase I, CACI, Inc. Federal Arlington, Virginia.
- Kaplan, I. T. (1980). Information flow in battalion command groups (ARI Technical Report 499). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A109 469)
- Kaplan, I. T. (1987). Lessons learned in research on command group training (ARI Technical Report 751). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A186 996)
- Kaplan, I. T. & Barber, H. F. (1979a). Evaluation of a computer-assisted battle simulation: CAMMS versus a CPX (ARI Technical Paper 355). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A068 014)
- Kaplan, I. T. & Barber, H. F. (1979b). Training battalion command groups in simulated combat: Identifications and measurement of critical performances (ARI Technical Paper 376). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A075 414)

- Krumm, R. L., Robins, J. E., & Ryan, T. G. (1973). Research on tactical decision making: III Predictor variables and criterion measures (ARI Technical Research Note 229). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A765 457)
- Lackey, L.L., DeLuca, A.J., & Tremble, J.R. (1975). Decision making and training techniques for command and control systems. Part 3. A model for combat order production (HumRRO Report FR-CD(C)-75-3). Alexandria, VA: Human Resources Research Organization.
- Metlay, W., Halatyn, A., & Liebling, D. (1983). Videotapes of a group planning process: An objective and quantitative method for analyzing behavior. Hofstra University, Department of Psychology, Applied Research Doctoral Program, Research Report Series.
- Metlay, W., Liebling, D., Silverstein, N., Halatyn, A., Zimberg, A., & Richter, F. (1985). Methodology for the assessment of the command group planning process. Hofstra University, Department of Psychology, Applied Research and Evaluation Doctoral Program, Research Report Series.
- Olmstead, J. A., Baranick, M. J., & Elder, B. L. (1978a). Training for brigade command groups: Training objectives and strategies (Technical Report TM-6008/001/00). Systems Development Corporation, Santa Monica, CA.
- Olmstead, J. A., Baranick, M. J., & Elder, B. L. (1978b). A training feedback system for brigade command groups (ARI Technical Report TR-78-A19). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A056 342)
- Olmstead, J. A., Baranick, M. J., & Elder, B. L. (1978c). Research on training for brigade command groups: Factors contributing to unit combat readiness (ARI Technical Report TR-78-A18). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A056 054)
- Olmstead, J. A., Christensen, H. E. (1973). Effects of work contexts: An intensive field study. Research Report No. 2, National Study of Social Welfare and Rehabilitation Workers, Work, and Organizational Contexts. Washington, D.C.: Department of Health, Education, and Welfare, Social and Rehabilitation Services.
- Olmstead, J. A., Christensen, H. E., & Lackey, L. L. (1973). Components of organizational competence: test of a conceptual framework (HumRRO TR-73-19). Alexandria, VA: Human Resources Research Organization.
- Olmstead, J. A., Christensen, H. E., Salter, J. A., & Lackey, L. L. (1975). Effects of work contexts in public welfare financial assistance agencies (Technical Report 75-7). Alexandria, VA: Human Resources Research Organization.

- Olmstead, J. A. Elder, B. L., & Forsyth, J. M. (1978). Organizational process and combat readiness: Feasibility of training organizational effectiveness staff officers to assess command group performance (ARI Technical Report 468). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A138 881)
- Otis, D. (1984). The effect of leadership on small group communication. Unpublished doctoral dissertation, Hofstra University, NY.
- Owens, J. D. & Brown, G. G. (1984). An investigation of the impact of headquarters structures on the military command environment. Naval Post-graduate School, Monterey California.
- Robins, J. E., Buffardi, L., & Ryan, T. G., (1974). Research on tactical decision making: Application of a decision prediction concept in a SIMTOS environment (ARI Technical Paper 246). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A780 812)
- Schein, E. H. (1965). Organizational psychology, Prentice-Hall, Inc., Englewood Cliffs, New Jersey.
- Shubik, M. (1975). Games for society, business, and war. American Elsevier Publishing Company, New York, NY.
- Solick, R. E., & Lussier, J. W. (1988). Design of battle simulations for command and control and staff training (ARI Technical Report 788). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A196 655)
- Strack, C. W. (1985). Command and control effectiveness experiments: Estimating the impact of C² network connectivity, and comparing geographic and functional C² organization. Defense Systems, Inc., McLean, Virginia.
- Sulzen, R. H. (1986). Annotated bibliography of tactical engagement simulation 1966-1984 (ARI Technical Report 725). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A178 671)
- Thomas, G. S. (1983). Battle simulation outcomes as potential measures of BCG performance in CATTS exercises (ARI Research Note). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.
- Thomas, G. S., & Cocklin, T. G. (1983). A model of mission accomplishment in simulated battle (ARI Technical Report 599). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (ADA 148 234)
- Thomas, G. S., Barber, H. F., & Kaplan, I. T. (1983). Improvement in command group performance as a result of diagnostic feedback and CATTS training (ARI Research Note). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

- Thomas, G. S., Barber, H. F., & Kaplan, I. T. (1984). The impact of CATTs system characteristics on selected measures of battalion command group performance (ARI Technical Report 609). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A140 231)
- Thomas, G. S., Kaplan, I. T., & Barber, H. F. (1984). Command and control training in the Combined Arms Tactical Training Simulator (ARI Technical Report 615). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A142 742)
- Tiede, R. V., & Leake, L. A. (1971). A method for evaluating the combat effectiveness of a tactical information system in a field army. Operations Research, Vol 19, pp 587-604.
- Tremble, T. R. Jr., DeLuca, A. J., & Lackey, R. (1975a). Decision-making and training techniques for command and control systems, Part 1. Overview of research (HumRRO Report FR-CD(C)-75-3). Alexandria, VA: Human Resources Research Organization.
- Tremble, T. R. Jr., DeLuca, A. J., & Lackey, R. (1975b). Decision-making and training techniques for command and control systems, Part 2. Terminal performance objectives for the demonstration model of CATTs (HumRRO Report FR-CD(C)-75-3). Alexandria, VA: Human Resources Research Organization.
- U.S. Army Reference Book (December, 1978). Guide to assessment of organizational process performance of battle staffs (RB 26-5). U.S. Army Organizational Training Center, Fort Ord, California.

APPENDIX A

PROCESS PERFORMANCE ASSESSMENT

SENSING

- Was all information available to the organization obtained by it?
- Were attempts to obtain information relevant and effective?
- Was correct interpretation placed upon information that was obtained?
- In view of the information available to the organization, was a correct assessment made?

COMMUNICATING INFORMATION

- Was information sensed by the organization communicated to everyone who needed it when they needed it?
- Was communication of information complete, accurate, and timely?
- Was communication of information efficient?

DECISION-MAKING

- Was all relevant, available information used in decision-making?
- Were decisions made at each level correct in view of information available to decision makers?

STABILIZING

- When decisions were made, were their potential effects upon the organization taken into account and action taken to counter any negative effects?
- Were internal operations or organizational arrangements adjusted appropriately to accommodate to new decisions?

COMMUNICATING IMPLEMENTATION

- After decisions, was communication about implementation requirements complete, accurate, and timely?
- Was everyone informed about implementation decisions and requirements that should be informed?

COPING ACTIONS

- Was execution of actions correct and effective?
- Were all actions leading from decisions actually carried out?

FEEDBACK

- Was action taken to obtain information about the outcomes of actions and decisions?
- Was information obtained in follow-up used to modify activities or make new decisions?

APPENDIX B

PROCESS DEFINITIONS AND CRITERIA

Process	Definition	Identification Criteria	Evaluation Criteria
Sensing	The act of acquiring information from or concerning any environment of the organization.	- Any act by a player of receiving, obtaining, or attempting to obtain information, orders, instructions, or recommendations from someone or something outside of the simulated organization. May involve passive receipt of information without initiative to obtain it, or may involve active attempt to obtain information. - Involves player-controller interaction in any mode. - In gaming simulations where company commander engages in board play as a "player/controller," Sensing includes company commander's "receipt" of information from the board, i.e., his sensing of board play.	- Accurate detection of all available information. - Correct interpretation of all detected information, to include appropriate weighting of its importance. - Accurate discrimination between relevant and irrelevant information. - Attempts to obtain information are relevant to mission, task, or problem. - Sensing activities are timely in relation to information requirements and the tactical situation.
Communicating Information	Those activities through which information which has been sensed by some representative of the organization is made available to those who must act on it or make decisions about it.	- Transmission and discussion of information by players after it has been sensed and before a decision has been made about it. - May pass through several links between sensing personnel and decision makers. - Player-player interaction, except where player informs Brigade controllers or subordinate unit controllers and information sensed. - May involve: - Initial transmittal of information by sensing individual. - Passing on of information by linking personnel. - Dissemination of information throughout organization. - Discussion and interpretation—Discussion for clarification or implication. - Includes communication of recommendations from subordinate units.	- Accuracy of transmission of available information. - Sufficiently complete to transmit full and adequate understanding to the receiver. - Timeliness appropriate to unit requirements. - Correct choice of recipients. - Whether message should have been communicated.

(Continued)

Process	Definition	Identification Criteria	Evaluative Criteria
Decision Making	Deliberative activities of one or more persons leading to a conclusion that some action will, or should, be taken by the organization. Usually evidenced by the initial communication of the decision by the decision maker. The communication may take the forms of announcement of the decision, a command, an order, or instructions. Decisions may lead to Active Sensing, formal Sensing Actions, Stabilizing Actions, Coping Actions, or Feedback Actions. Decision making includes decisions to rescind decisions. Decision making is not limited to commanders, it may include all players.	1. A communication of some sort reflecting the intention to take some kind of action. 2. Most often, the first evidence that a decision has been made will be a command, order, or instruction (oral or written) issued by the decision maker. 3. Usually player-player interaction; but at lower boundary of simulated organization, may involve player-controller interaction. 4. In Pegasus, the only evidence of a decision made by a company commander may be his interaction with the game board, i.e., his movement of pieces on the board (this would be a combination in one activity of decision making and coping action).	1. Adequacy—Was the decision adequately correct in view of circumstances and information available to the decision maker? 2. Appropriateness—Was the decision timely in view of the information available to the decision maker? 3. Completeness—Did the decision take into account all or most contingencies, alternatives, and possibilities?
Stabilizing	Actions intended to adjust internal operations, maintain internal stability or unit integrity, or prevent disruptions and negative side effects, as a consequence of coping actions. All actions intended to prevent potential negative effects to the organization which might occur because of Coping Actions.	1. Player/player interaction. 2. Limited to actions specifically intended to moderate the potential side effects of Coping Actions or to adjust internal organization or operations necessitated by the potential effects of a Coping Action.	1. Adequacy—Action is correct in view of the operational situation and the decision on order from which the action derived. 2. Appropriateness—Timing is appropriate in view of situation and the decision or order from which it derived. Choice of recipient of the action is appropriate. 3. Completeness—Action fully implements the decision from which it derived or fully meets the requirements of the situation.
Communicating Implementation	Those activities through which decisions and requirements resulting from decisions are communicated to those individuals or units who must implement the decisions. Includes (1) transmission of orders or instructions and (2) "discussion and interpretation—those communications through which clarification is achieved and implications for action are discussed. Includes all communication links between decision maker and final implementer of decision.	1. Player/player interaction. 2. Occurs after decision and before action. 3. Includes orders, instructions, and discussion of them and their implications, including clarification and attempts to obtain clarification. 4. Limited to communication about actions to be taken 5. May pass through several links between decision maker and executor of action.	1. Accuracy of transmission of instructions. 2. Sufficient completeness to transmit adequate and full understanding of actions required. 3. Timely transmission in view of both available information and the action requirements of participants. 4. Transmission to appropriate recipients. 5. Whether message should have been communicated.

(Continued)

Process	Definition	Identification Criteria	Evaluative Criteria
Coping Actions	The process of executing actions against target environments. Primarily concerned with execution and with how actions are carried out.	1. Player/controller interaction. 2. Actions taken at the point of contact with target environments, i.e., at boundaries of simulated organization. 3. Actions to "do something to" the external environment, i.e., to change or cope with the target environment. 4. Does <u>not</u> include actions to obtain information. 5. In battle simulations, coping actions may take form of orders or instructions to subordinate units played by controllers. 6. In battle simulations where company commander is a player/controller interacting with a game board, coping actions are his execution of the board play.	1. Correctness of action in view of both the operational circumstances and the decision or order from which the action derives. 2. Timeliness of the action in view of both operational circumstances and the decision or order from which the action derives. 3. Correctness of choice of target of the action. 4. Adequacy of execution of action.
Feedback	Activities that assist the organization to evaluate the effectiveness of its actions and that provide information upon which adjustments and future actions can be based.	1. Formal actions taken to obtain information about the results or effects of Coping Actions. 2. Player/controller interaction only. 3. Should be preceded by an organizational decision to initiate a feedback action.	1. Correctness of the action in view of both the operational circumstances and the decision or order from which the action derives. 2. Timeliness of the action in view of both the operational circumstances and the decision or order from which the action derives. 3. Correctness of choice of target of the action. 4. Adequacy of execution of the action.

APPENDIX C

Battalion Identification _____ Module No. _____
Type of Operation _____
Rater _____

SUMMARY PROCESS ASSESSMENT FORM

This form is to be completed at the end of each PEGASUS module. Use the scale shown below to rate the overall performance of the battalion players (battalion command group and company commanders) as a unit on each of the listed processes during the module. Use the scale to select the description that best fits your assessment of the battalion's performance of the process being rated. Enter in the space preceding each process, the number of the descriptor that best fits your assessment.

Scale

- 4 - Excellent
- 3 - Adequate
- 2 - Marginal
- 1 - Poor

<u>Rating</u>	<u>Process</u>
_____	Sensing
_____	Communicating Information
_____	Decision Making
_____	Stabilizing
_____	Communicating Implementation
_____	Coping Actions
_____	Feedback

APPENDIX D

Battalion _____

Controller Position _____

CONTROLLER RATING FORM

Please rate the battalion indicated above on each item below. For each item, make your judgment according to the combat outcome of the exercise. Do not rate the performance of the command group alone. Instead, rate the battalion, as a whole, in terms of its combat results. If several types of operations were conducted, e.g., attack, defense, etc., and the unit did better on some operations than others, base your ratings on your overall impression of all of the combat results achieved by the unit. For each item, place an "X" on the scale point that represents your best judgment.

1. Rate the extent to which the battalion accomplished its mission stipulated in the Brigade OPORD, including any changes subsequent to the initial OPORD.

1	2	3	4	5
Mission requirements were not met in any way		Unit was reasonably effective in view of circumstances		Mission was accomplished in all respects

2. Rate the extent to which area, or geographical objectives, was controlled by the battalion in accordance with its mission, to include any time schedules stipulated in its mission for holding areas or taking objectives.

1	2	3	4	5
Mission area requirements were not met in any way		Mission requirements were only partly met with respect to control of area		Control of area was in full compliance with mission requirements, including time schedules

3. Rate the expenditure of resources (personnel, weapons, equipment) by the battalion. (For this item, do not consider ratio of friendly to enemy losses. Rate only in terms of the condition of the battalion at the end of the exercise or operations).

1	2	3	4	5
Resources remaining make unit ineffective should be replaced		Resources remaining are not more than 75 nor less than 60 percent		Remaining strength such that operation can be continued without interruption.

4. Rate the expenditure of resources (personnel, weapons, equipment) by the battalion in relation to losses of resources by the threat forces opposing it, i.e., the ratio of friendly to threat losses of resources.

1	2	3	4	5
Loss ratio is totally in favor of threat forces		Losses were balanced equally between threat and friendly forces		Loss ratio is totally in favor of friendly battalion

5. Rate the overall combat effectiveness of this battalion in the exercise.

1	2	3	4	5
Totally ineffective		Minimally adequate for mission accomplishment		Effectiveness exceeded mission requirements

APPENDIX E

List of Terminal Performance Objectives (TPO) and Enabling Objectives (EO) for CATTs

TPO#1. Prevent unanticipated enemy actions by collecting information about the enemy, evaluating it with respect to the tactical situation, and using it effectively.

- EO1A. Collect information about the enemy.
- EO1B. Record information reported to the command group.
- EO1C. Evaluate information with respect to tactical situation.
- EO1D. Use available intelligence.

TPO#2. Recognize features in the area of operations (e.g., avenues of approach, key terrain features) and use their advantages and disadvantages to enhance the capabilities of friendly forces.

- EO2A. Identify avenues of approach into enemy's rear area during offensive operations.
- EO2B. Identify avenues of approach into battalion area of operations during defensive operations.
- EO2C. Identify key terrain features within the area of operations.
- EO2D. Utilize the advantages and disadvantages of key terrain and avenues of approach.

TPO#3. Determine the enemy's most probable course of action.

- EO3A. Know the enemies strength and location.
- EO3B. Know the enemies activities.

TPO#4. Identify and interpret the fires of enemy weapons.

- EO4A. Identify the fires of enemy weapons from various reports.
- EO4B. Interpret the fires of enemy weapons from various reports.

TPO#5. Identify enemy actions, based on available information, and take appropriate actions.

- EO5A. Identify anticipated enemy actions and take appropriate action.
- EO5B. Identify unanticipated enemy actions and take appropriate action.

TPO#6. Identify enemy feints, ruses, and deceptions.

No EOs specified.

TPO#7. Prevent unanticipated actions of the enemy by taking measures to counter the enemy's gathering of information about the situation of friendly forces.

E07A. Prevent unanticipated enemy actions by denying to the enemy knowledge of own activities through use of physical security measures.

E07B. Prevent unanticipated enemy actions by denying to the enemy knowledge of own activities through use of physical security measures.

E07C. Prevent unanticipated enemy actions by denying to the enemy knowledge gained through monitoring FM radio transmissions.

TPO#8. Employ organic and supporting fires to maximize their capabilities.

E08A. Employ organic indirect fires to their maximum destructive capabilities.

E08B. Employ supporting indirect fires to their maximum destructive capabilities.

E08C. Employ aviation fire support to its maximum destructive capability.

TPO#9. Implement improvisation when the conventional methods of employment of tactical formations, weapons, and equipment would not be effective.

E09A. Formulate tentative courses of actions and test their feasibility and effectiveness prior to implementation.

TPO#10. Identify, under conditions of unanticipated catastrophic equipment losses and casualties, remaining capabilities and vulnerabilities and take immediate action.

E010A. Formualte tentative course of action and test their feasibility and effectiveness prior to implementation.

TPO#11. Take effective actions in situations for which tactical doctrine is insufficient or inappropriate.

E011A. Formulate tentative actions and test their feasibility and effectiveness prior to implementation.

TPO#12. Effect timely resupply in personnel, supplies, and equipment to minimize interruptions in existing tactical operations.

E012A. Effect resupply of personnel.

E012B. Effect resupply of critical supplies.

E012C. Effect resupply of equipment.

TPO#13. Maintain continuous and secure communications.

E013A. Maintain continuous communications.

E013B. Maintain secure communications.

TPO#14. Communicate with subordinates and among members of the command group in a fashion that produces intended responses.

E014A. When communicating, use interpersonal knowledges and skills.

E014B. When communicating, use technical knowledge and skills.

TPO#15. During combined arms operations, alter the scheme of maneuver and/or plan of fire support when either or both would be ineffective due to unanticipated conditions associated with the enemy, weather, and terrain.

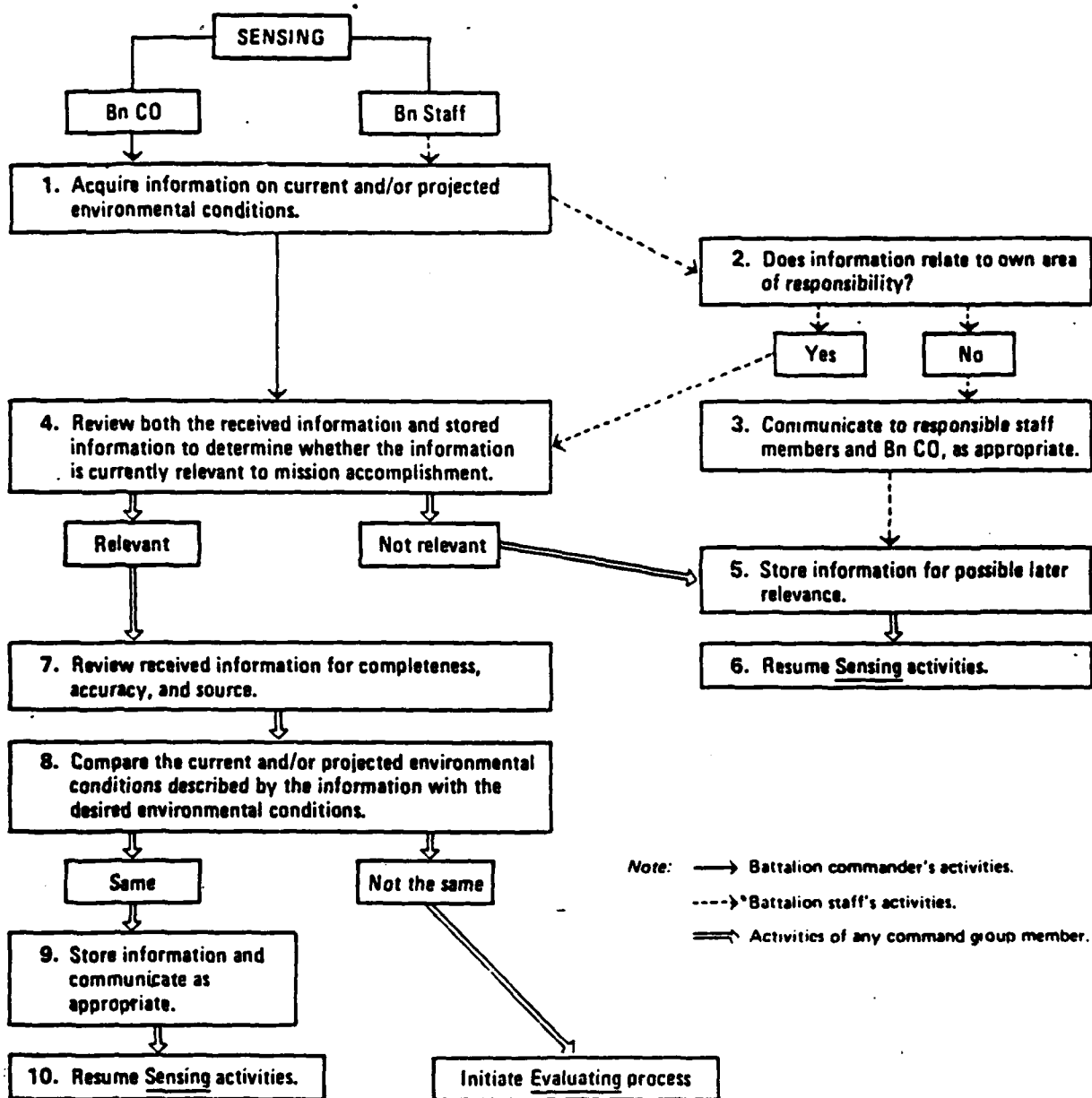
E015A. Change the task organization to alter the scheme of maneuver.

E015B. Change control measures to alter the scheme of maneuver.

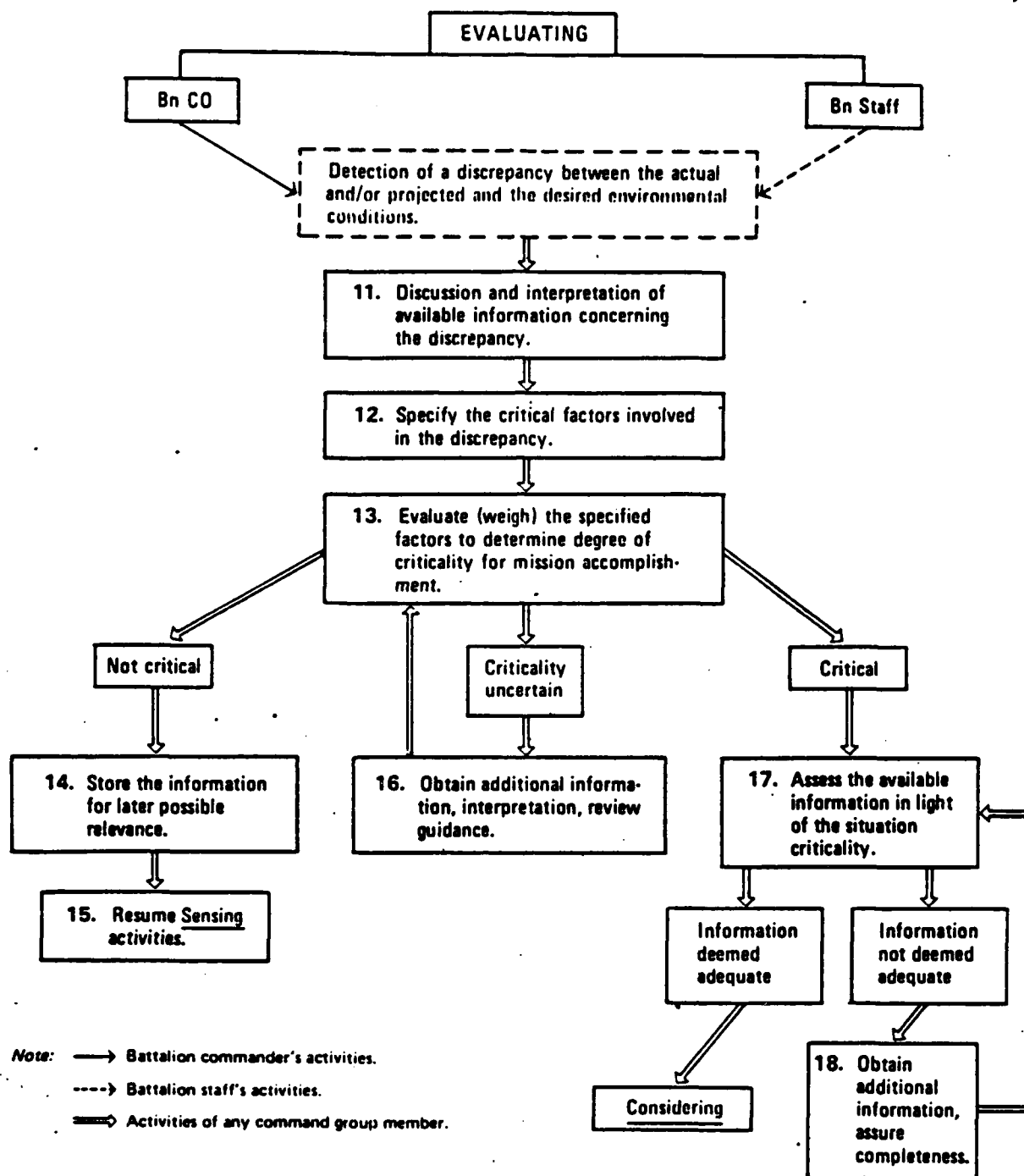
E015C. Change priorities of fires to alter the plan of fire support.

APPENDIX F

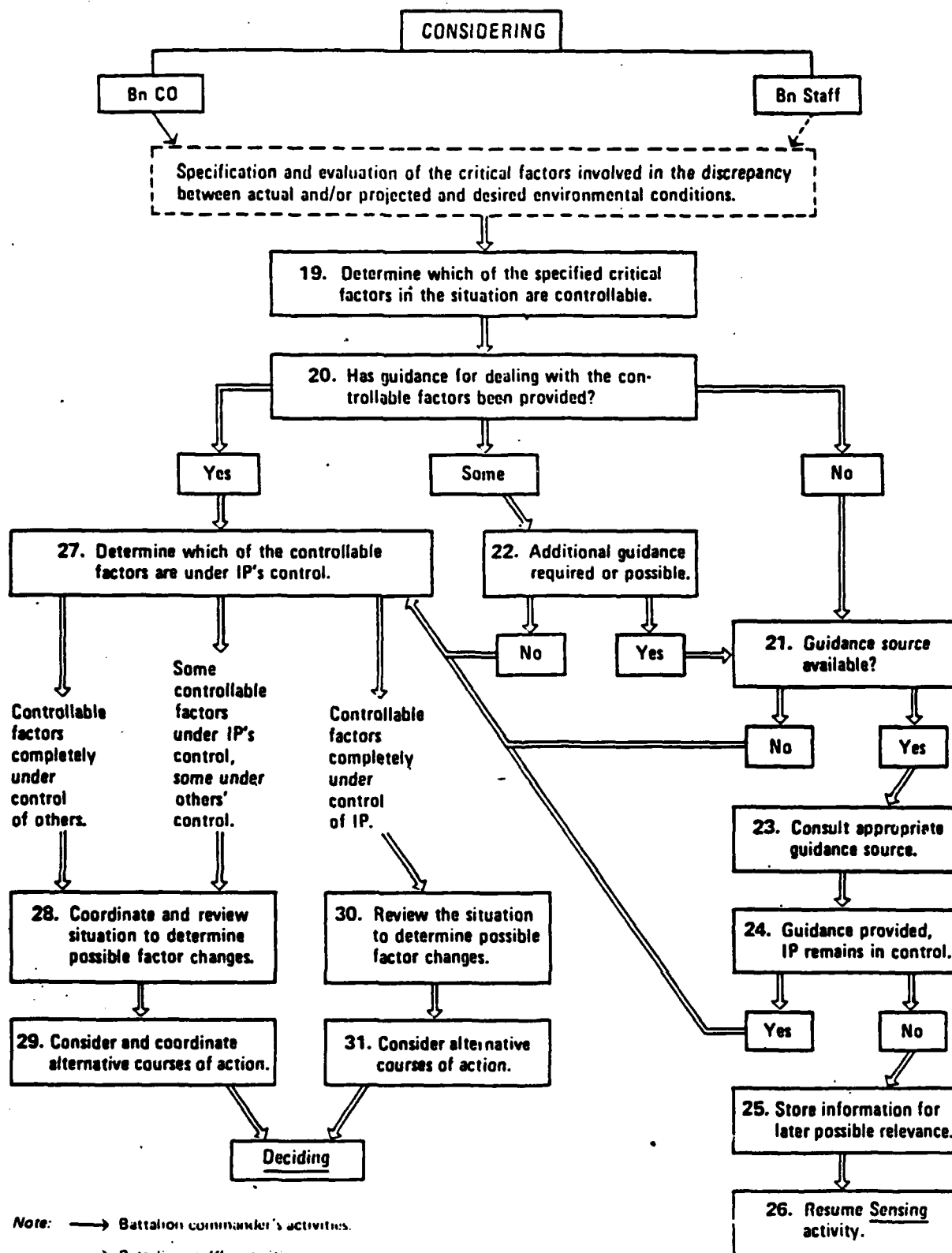
FLOW CHARTS OF THE CATEGORIES OF ACTIVITIES FOR EACH PROCESS



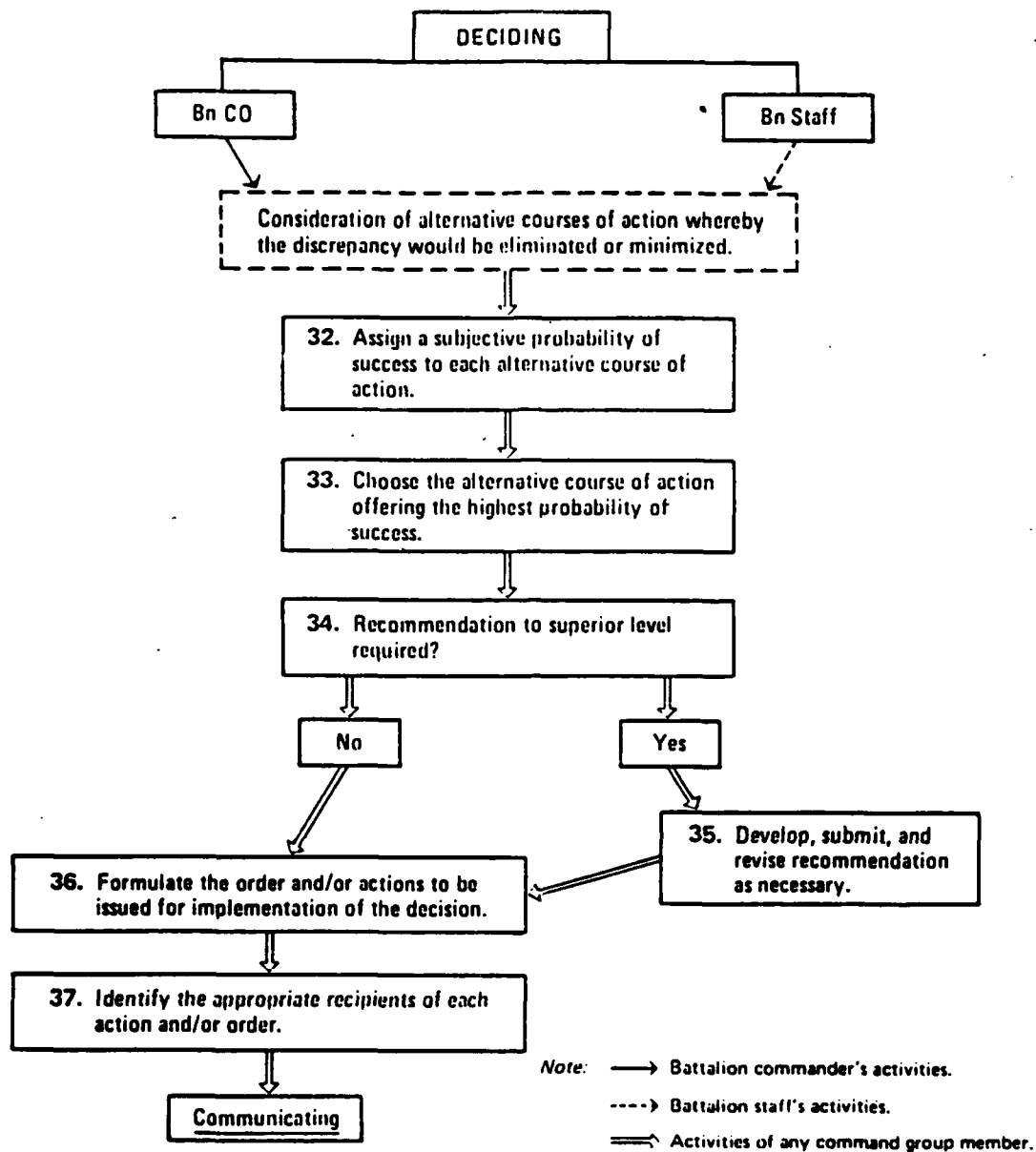
Categories of Activities in Sensing



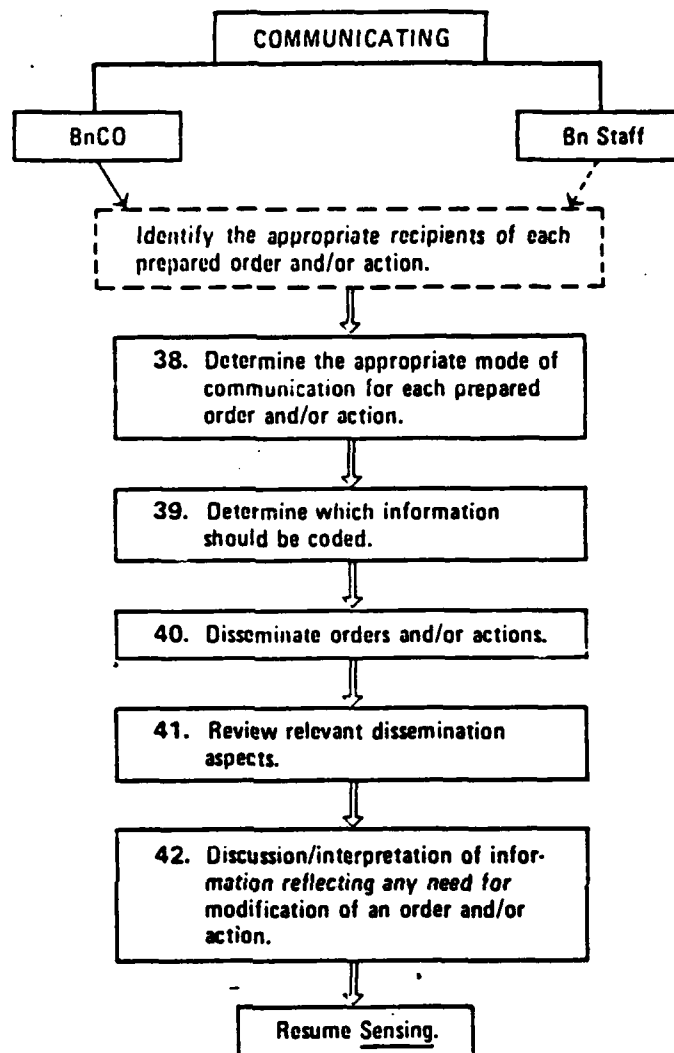
Categories of Activities in Evaluating



Categories of Activities in Considering



Categories of Activities in Deciding



Note: —→ Battalion commander's activities.
 ----→ Battalion staff's activities.
 ==→ Activities of any command group member.

Categories of Activities in Communicating

APPENDIX G

Information Items Evaluated

A. Situation

1. When the enemy situation is described, the briefer identifies and locates the position of the opposing force, describes its current and projected activities, outlines its current capabilities, and indicates its possible avenues of approach.

2. When the briefer discusses disposition of friendly forces, he describes the mission and status of higher headquarters, adjacent headquarters, and supporting units and also discusses attachments/detachments in terms of identification, location, and mission.

B. Mission

1. When the briefer describes the original mission, he states its overall purpose, its implementing objectives, the proposed approach to supporting the objectives, the element of the force charged with the mission, the location/terrain on which the task force will operate, and a timetable for mission execution.

2. In describing a related mission, the briefer provides the same information above and deviations from the original mission parameters.

C. Execution

1. When the briefer discusses his scheme of maneuver he should relate it to the mission, review enemy capabilities, review roles of subordinate units, assure proper task organization, and consider the terrain and weather.

2. When the briefer discusses the scheme of supporting fires, he should describe the support scheme of maneuver, preplanned fires, and priority of fires.

3. When the briefer assigns a mission element to a company task force or support section, he should describe the mission objectives, routes of withdrawal, actions at phase lines, and occupation of delaying positions.

4. When the briefer assigns the reserve responsibility to a company task force, he should describe the location to be occupied, preparations/priorities for commitment, routes of withdrawal, actions at phase lines and delaying position.

5. When the briefer gives coordination, instructions, he should caution the force not to conduct position defense, not to become decisively engaged to counterattack, only to disengage, to force enemy to deploy, to inflict maximum casualties, and to determine location of enemy's main effort.

D. Service Support

1. When describing service support effort, the briefer should discuss location, displacement, and security.
2. When discussing materials and services, the briefer should cover supply distribution, water points, 48 hour replacement time, transportation issues, and maintenance.
3. When discussing medivac and hospitalization, the location/displacement of the battalion and company and station should be described and provisions made for aeromedivac of urgent cases.
4. Personnel related issues to be discussed include the description of POW collecting points and planning of civilian/military operations.

E. Command and Signal

1. When the briefer discusses signal, he should describe the communication and electronics operating instructions, assignment of call signs and frequencies.
2. When discussing the command function, he should describe the TOC location and criteria for move on order. He should also identify the commander's location at the start of the engagement.

APPENDIX H

Critical Battalion Command Group Tasks, and Number of Subtasks, Described in ARTEP 71-2

- Task 1. Develop plan based on mission.
 - a. Analyze mission.
 - b.* Identify critical combat information and intelligence.
 - c. Identify critical friendly information.
 - d. Analyze friendly capabilities.
 - e. Select/control key terrain.
 - f. Select routes/zones to objective.
 - g. Select battle positions.
 - h. Select delay and covering force positions.
 - i.* Plan use of organic/attached and non-organic fires.
 - j. Determine priority of fires.
 - k. Determine fire support required.
 - l. Conduct initial fire support coordination.
- Task 2. Initiate intelligence preparation of the battlefield.
 - a.* Identify critical combat information and intelligence.
 - b.* Gather critical combat information and intelligence.
 - c.* Analyze opposing force.
 - d. Disseminate critical combat information and intelligence.
- Task 3. Prepare and organize the battlefield.
 - a. Determine critical place.
 - b. Select a course of action.
 - c. Organize for combat.
 - d. Select control measures.
 - e. Plan organic, attached, and non-organic supporting fires and determine priority.
 - f. Develop a communication plan.
 - g.* Communicate/coordinate plans and orders.
 - h. Reinforce terrain.
 - i. Plan/employ active/passive security measures.
 - j. Provide supplies.
 - k. Maintain equipment.
- Task 4. Troop lead-plan.
 - a. Supervise preparations.
 - b. Supervise compliance with TF order.
 - c. Conduct rehearsals.
- Task 5. See the battlefield during the battle.
 - a. Identify critical combat information and intelligence.
 - b.* Gather critical combat information and intelligence.
 - c.* Analyze opposing force.
 - d.* Disseminate critical combat information and intelligence.
- Task 6. Control and coordinate combat operations.
 - a. Modify scheme of maneuver.
 - b. Coordinate/communicate changes.
 - c. Supervise execution.
 - d. Maintain the battlefield.

- Task 7. Employ fires and other combat support assets.
 - a. Modify fire support plan.
 - b. Employ (to include organic/attached weapons systems and supporting artillery, air defense, TAC air, and attack helicopters).
 - c. Employ other combat support assets.
- Task 8. Concentrate/shift combat power.
 - a.* Determine critical place and time.
 - b.* Concentrate/shift combat power in the attack.
 - c.* Concentrate/shift combat power in the defense or retro-grade.
 - d. Protect thinly held areas.
- Task 9. Manage combat service support assets.
 - a. Arm and fuel the systems.
 - b. Fix the systems.
 - c. Support the troops.
 - d. Integrate CSS into scheme of maneuver.
- Task 10. Secure and protect the task force.
 - a.* Defeat or suppress opposing force's electromagnetic intelligence effort.
 - b. Defeat or suppress opposing force's imagery intelligence effort.
 - c. Defeat or suppress opposing force's human intelligence effort.
 - d. Deceive the opposing force.
 - e. Reduce vulnerability to opposing force mass destruction weapons systems.
 - f. Detect/impede threats to TF security.
 - g. Detect/defeat opposing force air assets.
- Task 11. Troop lead during battle.
 - a. Supervise compliance with TF order.
- Task 12. React to situations requiring special actions.
 - a.* React to opposing force electronic warfare.
 - b. React to chemical or biological attack.
 - c. React to nuclear attack.
 - d. React to loss of key member of command group.

* Subtasks that were weakly performed and significantly correlated with overall performance.

APPENDIX I

Performance Standards from Battalion ARTEP

PERFORMANCE STANDARDS

I. Gather and analyze required information.

A. Analyze mission.

Command group identifies specific and any implied tasks or constraints; addressing those tasks during the development of plan and in its own oral warning order/FRAGO/OPORD.

B. Determine what information is available and what additional information is required.

Command group determines information available on area of operations, friendly situation and opposing force situation.

Command group examines information on avenues of approach to objective; type, composition and location of TF and opposing maneuver and fire support units; TF and opposing force units capability of reinforcing by maneuver and fire; location of obstacles and opposing forces air attack and EW capability.

C. Determine what information sources are available.

Command group examines all appropriate resources. All assigned, attached, or DS units as well as higher echelon sources should be considered.

D. Gather all available information and request additional information as needed.

Command group gathers information on opposing force situation, areas of operations, and friendly situation. Records and displays are maintained and additional information requested as necessary from sources identified in Task 1C. As a minimum, collection efforts should focus on determining opposing force intentions and on determining status and situation of TF elements, major adjacent units and brigade reserve and supporting forces.

II. Develop a plan based on mission and modify it as required by events.

A. Determine friendly capabilities and limitations, request additional assets if needed.

Command group analyzes friendly capabilities (in terms of personnel, equipment and supplies, organic/attached/non-organic fires, maintenance and other supporting assets) to support current and contingency plans and requests additional assets if needed.

B. Estimate enemy capabilities and likely courses of action.

Command group, based on an understanding of known opposing force tactics and doctrine, compares that with combat information and intelligence received to determine opposing force intentions.

C. Identify key terrain.

Terrain which facilitates accomplishment of the TF mission is selected for control by occupation or fires. Terrain which, if captured/controlled by opposing forces, would facilitate accomplishment of the opposing force mission is designated key terrain.

D. Select battle position/routes to objectives.

Command group/staff selects appropriate battle positions/routes to objectives which reflects the commander's concept of accomplishment of the mission.

E. Identify critical place.

Command group determines the place on the battlefield where the TF combat power should be concentrated. Comparison made with OPFOR or chief controller's determination of critical place.

F. Develop and compare courses of action.

Based on the command group's analysis of the current situation and mission, courses of action are recommended and compared. The commander selects a course of action and clearly states his concept of the operation.

G. Individual staff planning: Commo.

Communications plan satisfies mission requirements, provides for COMSEC, specifies alternate means of communications, includes MIJL plan, and can be accomplished with the time and resources available to the TF.

H. Individual staff planning: Intel.

Plan provides for analysis of AO, intel estimates, intel requirements, minimizing TF vulnerability to mass destruction weapons, detecting impending threats to TF security and deceiving opposing force as to TF intentions.

I. Individual staff planning: Operations.

Command group task organizes the task force into company teams. A scheme of maneuver is developed to apply maximum combat power at the critical place while minimizing TF vulnerability. Operational security is addressed in plan.

J. Individual staff planning: Admin/Log.

Admin/log plan must complement scheme of maneuver. Priority of support (e.g., replacements, health services, classes of supply) is established. Planning must be flexible to allow changes during execution and facilitate future operations.

K. Individual staff planning: Fires.

Plan provides for preplanned fires, fires against targets of opportunity, suppression, surprise and deception, and air defense coverage while allowing TF elements to maneuver freely.

L. Coordinate with other staff members.

Command group selects control measures which support the scheme of maneuver, facilitate fire and movement by the TF and permit rapid changes as the battle develops. Command group examines components of the plan to identify areas that place limitations on, or require modification in, other components and resolve conflicts.

III. Communicate/coordinate.

A. Issue a warning order.

Upon receipt of a mission, warning order is issued to all necessary subordinate elements. Warning order includes nature of the operation and when and where TF OPORD will be issued. Communications and electronic security measures are rigidly adhered to throughout the TF.

B. Disseminate plans and orders.

Orders are coordinated with appropriate agencies. Orders are issued so as to allow TF elements maximum time for troop-leading procedures. Orders are appropriate, clear, and contain essential information. Changes in plans are communicated orally as a frag order and include changed objectives, control measures, and scheme of maneuver. Communications and electronic security measures are rigidly adhered to throughout the TF.

C. Disseminate combat information and intelligence.

Combat information should be event oriented rather than in periodic reports and summaries. Only information usable to the recipient should be disseminated. Information should be accurate and disseminated in time for the recipient to act upon it. Communications and electronic security measures are rigidly adhered to throughout the TF.

IV. Implement Plan

A. Concentrate/shift combat power.

Command group assesses the developing situation and issues directives to maximize use of combat power, requesting additional assets if necessary.

B. Reinforce terrain.

Command group tasks TF elements and supporting engineer unit to support the scheme of maneuver, reduces TF vulnerabilities, and increases OPFOR vulnerabilities (e.g., minefields and other obstacles, fortifications, etc.).

C. Provide supplies.

Coordinate with supporting supply elements to ensure that supplies (type and number), required to support TF elements and their weapon systems, are available to the TF on a timely basis.

D. Maintain equipment.

Command group determines status of equipment and directs repair/evaluation of non-operational equipment critical to mission accomplishment.

E. Request additional assets.

Command group supervises acquisition, control, and expeditious movement of replacements to points where they are needed. Command group requests reinforcements, if required, and supervises transfer of control.

V. Supervise combat operations.

A. Compare battlefield events with current order and concept of operations.

Command group monitors the developing situation, insuring that TF elements and supporting units comply with plans and orders, and that the friendly course of action continues to be appropriate.

B. Determine that a new course of action is necessary.

Command group detects a change in opposing force intentions, a threat to TF security or an inability to support their current course of action with available assets and determines that a change in plan is required.

C. Determine that a change in implementation is necessary.

Command group determines that some aspect of the plan are not being successfully implemented and makes necessary corrections.

APPENDIX J BRIGADE ARTEP SUBTASKS

Brigade ARTEP Subtasks found to be Usable in a CAMMS Based ARTEP, Organizational Process Scales, Overall Measures, and Controller Rating Form

PLANNING AND ORGANIZING PRIOR TO ENGAGEMENT

Planning Prior to Engagement

- 1A Analyze mission.
- 1E Select/control key terrain.
- 1F Select routes/zones of approach to objective.
- 1G Assign areas and sectors of defense/battle positions.
- 1H Select covering force positions.
- 3A Determine critical place.
- 3B Select a course of action.

Intelligence Activities Prior to Engagement

- 1B Identify critical combat information and intelligence.
- 1C Identify critical friendly information.
- 1D Analyze friendly capabilities.
- 2A Identify critical intelligence.
- 2B Gather critical information and intelligence.
- 2C Analyze enemy.
- 2D Disseminate critical combat information and intelligence.

Organizing the Battlefield Prior to Engagement

- 3C Organize for combat.
- 3D Select control measures.
- 3F Develop communication plan.
- 3G Communicate/coordinate plans and orders.
- 3J Provide supplies.

Planning Fire Support Prior to Engagement

- 1I Plan use of organic/attached and non/organic fires.
- 1J Determine priority of fires.
- 1L Conduct initial fire support coordination.
- 3E Continue to plan organic, attached, and non-organic supporting fires and determine priority.

FIGHTING THE BATTLE

Intelligence Activities During the Battle

- 5A Identify critical intelligence.
- 5B Gather critical combat information and intelligence.
- 5C Analyze enemy.
- 5D Disseminate critical combat information and intelligence.

Controlling and Coordinating Combat Operations During the Battle

- 6A Modify scheme of maneuver
- 6B Coordinate/communicate changes.
- 8A Determine critical place and time.
- 8C Concentrate/shift combat power in the defense/retrograde.
- 8D Protect thinly held areas.

Employing Fires and Fire Support Assets During the Battle

- 7C Employ other combat support assets.
- 9A Arm and fuel the system.
- 9C Support the troops.
- 9D Integrate CSS into scheme of maneuver.

Securing and Protecting the Task Force During the Battle

10A Defeat or suppress enemy's EM intelligence effort.

10E Reduce vulnerability to enemy mass reduction weapons system.

Reacting to Situation Requiring Special Actions During the Battle

12A React to enemy EW warfare.

12D React to loss of key member of command group.

_____ Planning. (101)	Overall, planning of the command group is: 1. Complete 2. Efficient 3. Covers all contingencies
_____ Decision Making. (102)	Overall, decision making of the command group is: 1. Timely 2. Correct in view of the situation
_____ Implementation. (103)	Overall, implementation of decisions is characterized by: 1. General supervision 2. Delegation of appropriate responsibilities 3. Timely and appropriate follow-up
_____ Communication. (104)	Overall, communication of the command group, both upward and downward, is: 1. Timely 2. Complete 3. Accurate 4. Efficient
_____ Responsiveness to Subordinate Units. (105)	Overall, the command group responds to requests and requirements of subordinate units: 1. Promptly 2. Helpfully 3. Accurately

- _____ Overall, how effective was this brigade command group? (106)
- _____ Overall, how effectively were S3 activities performed? (107)
- _____ Overall, how effectively were S2 activities performed? (108)
- _____ Overall, how effectively were S1 activities performed? (109)
- _____ Overall, how effectively were S4 activities performed? (110)

- NOTICE -

INFORMATION PROVIDED IN THIS QUESTIONNAIRE WILL NOT BE USED
IN ANY WAY TO EVALUATE YOUR UNIT

CONTROLLER RATING FORM

Instructions

The following pages list a number of critical activities performed by brigade command groups during the planning and execution phases of combat operations. For each activity, a standard of performance is also presented. Note that the activities are grouped into two phases -- (1) Planning and Organizing Prior to the Engagement, and (2) Fighting the Battle.

For each activity, please rate the performance of the brigade command group according to the standards presented. Note that you should rate the performance of the Brigade Command Group only. The activities of Battalion Command group players should not be a consideration in your ratings.

For each activity, enter in the space provided one of the numbers shown below which best fits your judgment of how well the brigade command group performed relative to the standard. Use the following scale to make your ratings.

Rating

Definition

- 0 ---- No personal knowledge of command group performance of this activity.
- 1 ---- Totally Ineffective. Activity was not performed or it included major deficiencies so that it was never completed.
- 2 ---- The command group's performance included several major deficiencies so that the activity, although completed, was ineffective.
- 3 ---- The command group's performance included many minor deficiencies so that, overall, the activity was marginal for mission accomplishment.
- 4 ---- Adequate. Command group performance was minimally adequate for mission accomplishment.
- 5 ---- Command group performance was better than adequate but included minor deficiencies.
- 6 ---- The quality of performance somewhat exceeded that required for mission accomplishment.
- 7 ---- Superior. Command group performance of this activity was complete in all respects and the quality of performance fully exceeded that required for mission accomplishment.

ARI Document Control
No. 77:5162B

Appendix K

Terminal Learning Objectives and Learning Objectives for Brigade Command Groups

Two types of objectives are presented below. Each Terminal Learning Objective, designated by a roman numeral, is shown first, followed immediately by the Learning Objective associated with it. Learning Objectives are designated by arabic numerals. Since each type of objective can be, at times, used or listed independently of the other types, each type is listed consecutively throughout their respective number series.

I. Given a simulated combat environment, an enemy force within the brigade tactical intelligence zone, and a division OPORD which includes enemy and friendly situations and a brigade mission, brigade command group will issue a warning order in sufficient time to permit effective planning and troop leading by subordinate units and a brigade OPORD which, commensurate with provided and available information:

- a. States mission of the brigade and each subordinate unit.
- b. Specifies the course of action which, in light of known information, offers the greatest probability of success.
- c. Establishes the task organization which best provides for execution of the selected course of action.
- d. Specifies the security measures required.
- e. Specifies tactical control measures required in addition to those specified by higher headquarters.
- f. Provides an obstacle/anti-armor plan, if appropriate, which effectively supports the selected course of action.
- g. Includes a communication plan which, commensurate with mission requirements, provides for COMSEC, specifies alternative means of communication, and insures effective operation of MLJI plan.
- h. Includes a fire support plan which provides for supporting preplanned fires, fires against targets of opportunity, suppressions, surprise deceptions, and air defense coverage.
- i. Is updated, modified, and communicated to subordinate units as required for effectiveness by receipt of additional information and intelligence.

1. Given a simulated combat environment, an enemy force within the brigade tactical intelligence zone, and a division OPORD which includes enemy and friendly situations, and a brigade OPORD which includes the following:

- (a) Accurate statement of brigade mission.
- (b) Statement of command relationships commensurate with mission.
- (c) Identification of all supporting friendly units and their missions.
- (d) Statement of the concept of operation, to include passage of lines, specific objectives for units, and required coordinations.

2. Given a simulated combat environment, an enemy force within the brigade tactical intelligence zone, and a division OPORD which includes enemy and friendly situations and a brigade mission, brigade command group will:

- (a) Identify terrain which will facilitate accomplishment of the brigade mission and terrain which will facilitate accomplishment of the enemy's mission.
- (b) Specify how key terrain should be controlled by fire or occupation, using brigade and available supporting resources.

3. Given a simulated combat environment, an enemy force within the brigade tactical intelligence zone, and a division OPORD which includes enemy and friendly situations and a brigade attack mission, brigade command group will specify avenues of approach which optimize the following conditions:

- (a) Provide maximum cover and concealment.
- (b) Minimize the effects of obstacles.
- (c) Permit mutual support and overwatch.
- (d) Permit effective employment of weapons.
- (e) Facilitate control while permitting units to maneuver.
- (f) Maximize brigade and unit mobility.
- (g) Capitalize on enemy vulnerabilities.
- (h) Minimize time for units to close on objective.
- (i) Facilitate logistical operations.

4. Given a simulated combat environment, an enemy force within the brigade tactical intelligence zone, and a division OPORD which includes enemy and friendly situations and a brigade defense mission, brigade command group will assign areas and sectors/battle positions which:

- (a) Block most critical avenue of approach into the defensive sector.
- (b) Minimize vulnerabilities to enemy's frontal direct fire weapons and indirect fire weapons.
- (c) Maximize capabilities of own weapons and permit engagement of targets at maximum effective range.
- (d) Exploit and reinforce natural terrain obstacles.
- (e) Permit mutual support and overwatch.
- (f) Facilitate control while permitting subordinates to deploy and maneuver.
- (g) Maximize Bde and Bn mobility by allowing for strong, quick counterattack.
- (h) Capitalize on enemy vulnerabilities.
- (i) Reduce vulnerability to air attacks.

5. Given a simulated combat environment, an enemy force within the brigade tactical intelligence zone, and a division OPORD which includes enemy and friendly situations and a brigade defense mission, brigade command group will identify initial and successive covering force positions which:

- (a) Block obvious critical avenues of approach.
- (b) Force enemy to deploy and concentrate forces repeatedly.
- (c) Maximize effectiveness of own weapons.
- (d) Force enemy to travel along exposed approach routes.
- (e) Use natural terrain/man made obstacles.
- (f) Use battle positions that facilitate transition to limited attack, defense, or withdrawal.

6. Given a simulated combat environment, an enemy force within the brigade tactical intelligence zone, and a division OPORD which includes enemy and friendly situations, an intelligence estimate, and a brigade mission, brigade command group will:

- (a) Identify the place on the battlefield where combat power should be concentrated.
- (b) Weight the combat power of the brigade elements in accordance with indications of the intelligence estimate.

7. Given a simulated combat environment, an enemy force within the brigade tactical intelligence zone, and a division OPORD which includes enemy and friendly situations, an intelligence estimate, and a brigade mission, brigade command group will develop a course of action which takes into account all significant factors in the information provided.

8. Given a simulated combat environment, an enemy force within the brigade tactical intelligence zone, and a division OPORD which includes enemy and friendly situations and a brigade mission, brigade command group will issue a brigade OPORD which:

- (a) Organizes the battalions into a combined arms task force best suited to the brigade mission.
- (b) Precisely states the task organization.
- (c) States missions of all organic maneuver elements.
- (d) States missions of all attached maneuver elements, if any.
- (e) Clearly states a scheme of maneuver capable of accomplishing the mission.

9. Given a simulated combat environment, an enemy force within the brigade tactical intelligence zone, a division OPORD which includes enemy and friendly situations and a brigade mission, and a brigade scheme of maneuver, brigade command group will specify control measures which:

- (a) Support the scheme of maneuver.
- (b) Facilitate fire and movement.
- (c) Permit rapid changes as battle develops.
- (d) Are explicitly defined.
- (e) Associated with recognizable terrain.
- (f) Sufficient to effectively support the mission.

10. Given a simulated combat environment, an enemy force within the brigade tactical intelligence zone, a division OPORD which includes enemy and friendly situations and a brigade mission, and a brigade scheme of maneuver, brigade command group will develop a communications plan which:

- (a) Effectively provides for COMSEC.
- (b) Specifies appropriate alternative means of communication.
- (c) Provides instructions or actions to be performed if jamming occurs.

11. Given a simulated combat environment, an enemy force within the brigade tactical intelligence zone, a division OPORD which includes enemy and friendly situations and a brigade mission, and a brigade OPORD, brigade command group will develop a fire support plan which:

- (a) Effectively provides supporting preplanned fires, fires against targets of opportunity, suppressions, surprise, deception, and air defense coverage.
- (b) Accurately identifies FA, 4.2, and AD priority of fires.
- (c) Specifically identifies fire support/target acquisition assets by type and availability.
- (d) Identifies specific fire support coordination measures.

11. Given a simulated combat environment, an enemy force within the brigade tactical intelligence zone, a division OPFOR which includes enemy and friendly situations and a brigade mission, and an incomplete intelligence summary, brigade command group will initiate information and intelligence gathering activities which:

- a. Reflecting accurate identification of intelligence shortfalls.
- b. Use all potential sources of information.
- c. Indicate awareness of significant tactical indicators, targets, and EW, NBC, and CAS capabilities of enemy forces.
- d. Effectively employ organic agencies to obtain all enemy forces.
- e. Effectively employ outside agencies to obtain all available information and intelligence.

12. Given a simulated combat environment, an enemy force within the brigade tactical intelligence zone, a division OPORD which includes enemy and friendly situations and a brigade mission, and an incomplete intelligence summary, brigade command group will identify the following:

- (a) Avenues of approach into defended areas.
- (b) Composition and size of enemy forces.

- (c) Enemy's scheme of maneuver and fire support.
- (d) Enemy's ability to attack by air.
- (e) Enemy's EW capability.

13. Given a simulated combat environment, an enemy force within the brigade tactical intelligence zone, a division OPORD which includes enemy and friendly situations and a brigade mission, and an incomplete intelligence summary, brigade command group will identify the following:

- (a) All friendly TF elements.
- (b) All major adjacent units.
- (c) The division reserve.
- (d) Major supporting forces.

14. Given a simulated combat environment, an enemy force within the brigade tactical intelligence zone, a division OPORD which includes enemy and friendly situations and a brigade mission, and an incomplete intelligence summary, brigade command group will analyze enemy capabilities and weaknesses and, within the information provided, accurately identify:

- (a) Location of enemy forces, to include reserve.
- (b) Time/distance factors for enemy to be reinforced.
- (c) Enemy weaknesses.

15. Given a simulated combat environment, an enemy force within the brigade tactical intelligence zone, a division OPORD which includes enemy and friendly situations and a brigade mission, and an incomplete intelligence summary, brigade command group will:

- (a) Correctly identify all combat information and intelligence shortfalls.
- (b) Specify required information and intelligence collection activities, to include: (1) effective utilization of all GSR elements, (2) deployment of personnel sufficiently to observe enemy prior to engagement, (3) effective employment of all organic agencies, and (4) effective employment of all available outside agencies.

III. Given a simulated combat environment and using division intelligence summaries, additional information provided by or obtained from division and outside agencies, and additional information provided by or obtained from organic sources, brigade command group will disseminate and continuously update intelligence to subordinate and adjacent units and higher headquarters, as appropriate, which meets the following requirements:

- a. Accurate
- b. Complete
- c. Timely
- d. Event oriented
- e. Relevant to recipient(s)

f. Subordinate elements provided with accurate estimates of probable enemy capabilities, actions, and intentions.

16. Given a simulated combat environment and using division intelligence summaries, additional information provided by or obtained from division and outside agencies, and additional information provided by or obtained from organic sources, brigade command group will prepare an intelligence estimate which, within available information, accurately analyzes enemy capabilities and estimates probable intentions and actions.

17. Given a simulated combat environment and using division intelligence summaries, additional information obtained from or provided by division and outside agencies, and additional information obtained from or provide by organic sources, brigade command group will disseminate and continuously update intelligence to subordinate and adjacent units and higher headquarters, as appropriate, which is: (1) accurate, (2) complete, (3) timely, (4) event oriented, (5) relevant to recipients, and (6) for subordinate units, provides them with accurate estimates of enemy capabilities, intentions, and actions.

IV. Given a simulated combat environment and a previously issued brigade OPORD, brigade command group will coordinate and supervise preparation for the battle by activities which:

a. Fully integrate and coordinate activities of members of the command group.

b. Monitor and coordinate preparation of subordinate elements in accordance with the OPORD.

c. Establish with supporting supply elements coordination that will insure that adequate supplies are available and capable of being issued as required to accomplish the mission.

d. Establish with all supporting units and agencies coordination that will insure that support will be available according to plan and as required.

V. Given a simulated combat environment, during engagement of brigade elements with enemy forces, brigade command group will maintain on a continuing basis information and intelligence collection activities which:

a. Task subordinate units to assist in determining tactical indicators and targets.

b. Request all appropriate sources to supply event oriented information on a continuing basis.

c. Request and receive reports of enemy activity from units in contact as engagement develops.

18. Given a simulated combat environment and a developing engagement of brigade elements with enemy forces, brigade command group will:

(a) Accurately identify EEI's.

(b) Task subordinate units to assist in determining tactical indicators and targets.

19. Given a simulated combat environment and a developing engagement of brigade elements with enemy forces, brigade command group will:

(a) On a continuing basis, accurately identify combat information and intelligence shortfalls.

(b) Aggressively gather required information from all available appropriate sources.

20. Given a simulated combat environment, a developing engagement of brigade elements with enemy forces, and continuing receipt of combat information and intelligence, brigade command group will:

(a) Periodically issue updated estimates which accurately predict enemy intentions in accordance with available information and intelligence.

VI. Given a simulated combat environment, during engagement of brigade elements with enemy forces, brigade command group will, on a continuing basis, disseminate and update to subordinate and adjacent units and higher headquarters, as appropriate, information and intelligence which is:

a. Accurate

b. Complete

- c. Timely
- d. Event oriented
- e. Relevant to recipient(s)

21. Given a simulated combat environment, an ongoing engagement of brigade elements with enemy forces, and periodic information an intelligence from a variety of sources, brigade command group will:

- (a) Accurately identify essential elements of intelligence which should be disseminated, to include recipient(s).
- (b) Issue appropriate messages which are correct in format, procedure, and content.

VII. Given a simulated combat environment, during engagement of brigade elements with enemy forces, brigade command group will control and coordinate combat operations. Control and coordination activities will include at a minimum:

a. Modifications of the scheme of maneuver which maximize cover, concealment, suppression, and teamwork, and exploit unanticipated changes in combat situation.

b. Redirection of subordinate unit activities which maximally employ combat power to exploit enemy weaknesses, counter unanticipated contingencies, or destroy the enemy force.

c. Approvals, modifications, or disapprovals of subordinate unit activities which are in accordance with the current or revised scheme of maneuver.

d. All essential coordination, including coordination with higher headquarters and adjacent units, is prompt and timely.

e. Communications of required changes to subordinate units are in the form of oral fragmentary orders which include changed objectives, control measures, and modified schemes of maneuvers.

f. Activities which will insure that LOC and vital logistical facilities remain open and usable and that barriers/minefields are continuously upgraded, improved, and maintained.

22. Given a simulated combat environment, an ongoing engagement of brigade elements with enemy forces, and a changing tactical situation, brigade command group will develop a modification to the scheme of maneuver which:

- (a) Stipulates a specific course of action.

- (b) Effectively meets mission requirements.
- (c) Effectively meets requirements of the changed tactical situation.
- (d) Emphasizes cover, concealment, suppression, and teamwork.
- (e) Effectively uses combat power to exploit enemy weaknesses at the critical time and place.

23. Given a simulated combat environment, an ongoing engagement of brigade elements with enemy forces, and a modified scheme of maneuver, brigade command group will specify all essential coordinations and changes to be communicated. Coordinations and change communications will be:

- (a) Accurate
- (b) Complete
- (c) For subordinate units, oral fragmentary orders which include changed objectives, control measures, and scheme of maneuver.

24. Given a simulated combat environment, an ongoing engagement of brigade elements with enemy forces, and a defense or retrograde situation, brigade command group will specify required concentration of shifts in its combat power according to weapons capabilities and movement of the enemy force, to include:

- (a) Effective redeployment of forces after enemy commitment.
- (b) Effective use of Air Force, field artillery, scouts, TOW's, smoke, and CSC.
- (c) Requests for additional units from division reserve, if required.
- (d) Effective organization and assignment of reinforcements.
- (e) Effective use of attached units such as ADA and Engineers.
- (f) Effective protection of thinly held areas.

VIII. Given a simulated combat environment, during engagement of brigade elements with enemy forces, brigade command group will direct employment of fires and other combat service support assets to:

- a. Effectively concentrate combat power at the most critical time and place.
- b. Effectively suppress enemy weapon systems which would interfere with accomplishment of the brigade mission.

c. Provide to the brigade all additional required resources.

25. Given a simulated combat environment, ongoing engagement of brigade elements with enemy forces, a changing tactical situation and fire support plan, brigade command group will modify the fire support plan, to include:

- (a) Effectively meet fire support requirements of the changed tactical situation.
- (b) Effective communication of modified plan to all brigade elements.
- (c) Effective coordination of changes with adjacent units and other concerned agencies.

26. Given a simulated combat environment, ongoing engagement of brigade elements with enemy forces, a changing tactical situation and a fire support plan, brigade command group will employ fire support plan, brigade command group will employ fire support effectively by:

- (a) Assigning requests for immediate fire support to appropriate support agencies.
- (b) Engagement of targets with appropriate weapons.
- (c) Engagement of targets at maximum range.
- (d) Concentration of fire support assists at critical place and time.
- (e) Effective suppression of enemy fires.
- (f) Monitoring of ammunition expenditures.

27. Given a simulated combat environment, ongoing engagement of brigade elements with enemy forces, and a changing tactical situation, brigade command group will task supporting combat engineers to:

- (a) Create obstacles and minefields.
- (b) Produce protective shelters.
- (c) Maintain tactical and supply routes.

IX. Given a simulated combat environment, ongoing engagement of brigade elements with enemy forces, and a changing tactical situation, brigade command group will manage combat service support assets to:

a. Make replacement personnel, POL, equipment and other critical supplies immediately available to and efficiently used by brigade elements.

b. Maintain maximum feasible percentages of functioning personnel, weapons, and equipment in the brigade.

28. Given a simulated combat environment, ongoing engagement of brigade elements with enemy forces, and a changing tactical situation, brigade command group will:

- (a) Handle expeditiously requests for supplies and equipment in accordance with SOP.
- (b) Insure that delivery of supplies to brigade elements is made as far forward as prudent.
- (c) Insure that troop subsistence meets requirements.
- (d) Insure movement of replacements to points where they are most required.

29. Given a simulated combat environment, ongoing engagement of brigade elements with enemy forces, and a changing tactical situation, brigade command group will maneuver CSS elements to:

- (a) Keep CSS elements in close proximity to weapons systems.

X. Given a simulated combat environment, ongoing engagement of brigade elements with enemy forces, and a changing tactical situation, brigade command group will employ measures to secure and protect the task force which prevent the enemy force from accurately determining brigade strength, task organization, force dispositions, vulnerabilities, capabilities, and intentions.

30. Given a simulated combat environment, ongoing engagement of brigade elements with enemy forces, and a changing tactical situation, brigade command group will employ communications and electronic security discipline which:

- (a) Prevents security violations during radio communication.
- (b) Prevents extraneous radio traffic not required by the tactical situation.
- (c) Employ approved radio procedures.

31. Given a simulated combat environment, ongoing engagement of brigade elements with enemy forces, and a changing tactical situation, brigade command group will reduce vulnerability to enemy mass reduction weapons by dispersing brigade elements to the extent allowed by the terrain and tactical situation.

XI. Given a simulated combat environment, ongoing engagement of brigade elements with enemy forces, a changing tactical situation, and the occurrence of situations which require special actions, brigade command group will:

- a. Continue operations with a minimum of confusion and disruption.
- b. Initiate specific measures which counter or alleviate the special situation.

32. Given a simulated environment, ongoing engagement of brigade elements with enemy forces, a changing tactical situation, and the occurrence of enemy EW activities, brigade command group will recognize enemy jamming activities and react to the situation by:

- (a) Continuing the operation on an appropriate routine basis.
- (b) Directing a switch to alternate frequency using proper authentication procedures.
- (c) Using a secure means of communication, submit a MIJI report to higher headquarters.

33. Given a simulated combat environment, ongoing engagement of brigade elements with enemy forces, a changing tactical situation, and the simulated loss of a key member of the command group, brigade command group will react to the situation by:

- (a) Continuing the operation without disruption or confusion.
- (b) Information division of the loss.
- (c) Requesting division to supply a replacement.
- (d) A member of the command group assuming duties of the lost personnel until a replacement arrives.

APPENDIX L

UNIT RATING FORM - CAMMS EXERCISE

- NOTICE -

**INFORMATION PROVIDED IN THIS QUESTIONNAIRE WILL NOT
BE USED IN ANY WAY TO EVALUATE YOUR UNIT**

Instructions

On the following pages appear 15 items on which you are requested to rate several aspects of your just completed CAMMS exercise. Use the rating scales that are provided to indicate your best judgment for each item.

The ratings are not to be used in any way as an evaluation of the unit participating in the exercise. The data that are provided by your ratings will be used to improve the training provided by CAMMS and contribute to better understanding of unit training requirements.

Please place an "X" in the blank that describes your role in the recently completed exercise.

- ☐ Brigade-level player
- ☐ Battalion-level player
- ☐ Company-level player
- ☐ Subordinate unit controller (friendly)
- ☐ Other controller.

ARI Document Control
No. 77:5162C

For the following items, use the scale shown below each item to indicate your rating of the item.
Enter the number for your rating in the space preceding the item.

Ratings

_____ 1. How was the planning during the planning phase in this exercise?

Totally ineffective		Incomplete or some major deficiencies		Complete, efficient, all contingencies covered		
1	2	3	4	5	6	7

_____ 2. How was the decision making in this exercise?

Many incorrect or untimely decisions		Occasionally untimely or incorrect		Timely and correct		
1	2	3	4	5	6	7

_____ 3. In general, how was the communication in the exercise?

Consistently incomplete, inaccurate, and untimely		Sometimes incomplete, inaccurate, and untimely		Timely complete, accurate, efficient		
1	2	3	4	5	6	7

_____ 4. How was the responsiveness of your superior levels to the requests and requirements of subordinate units?

Totally unresponsive			Occasionally unresponsive		Prompt, helpful, correct	
1	2	3	4	5	6	7

_____ 5. How were the intelligence collection activities in this exercise?

Late incomplete, and inaccurate			Sometimes aggressive but sometimes passive		Thorough and aggressive	
1	2	3	4	5	6	7

_____ 6. How was the provision of intelligence and information to you in this exercise?

Late, incomplete and inaccurate			Occasionally late, incomplete, or inaccurate		Timely, complete accurate	
1	2	3	4	5	6	7

_____ 7. How was the coordination between adjacent units in this exercise?

No coordination occurred			Occasional breakdowns occurred		Frequent and effective	
1	2	3	4	5	6	7

_____ 8. How good a job was done in anticipating contingencies that might occur during the engagement and preventing them from occurring or minimizing their effects?

Many contingencies overlooked

Some contingencies overlooked

Most contingencies anticipated and prevented

1 2 3 4 5 6 7

_____ 9. When adjustments or changes in plans or operations were necessary, how timely were they?

Never timely

Sometimes timely, sometimes not timely

Always timely

1 2 3 4 5 6 7

_____ 10. When adjustments or changes in plans or operations were necessary, how effectively were they accomplished?

Consistently ineffective

Sometimes effective, sometimes not effective

Consistently effective

1 2 3 4 5 6 7

_____ 11. How clear were you about the mission objectives of the brigade and battalion?

Never clear at any time

Fairly clear most of the time

Perfectly clear all the time

1 2 3 4 5 6 7

_____ 12. How clear were you about what was expected of you by your commander during the exercise?

Never clear
at any time

Fairly clear
most of the time

Perfectly clear
at all times

1 2 3 4 5 6 7

_____ 13. Rate the amount of supervision exercised over your unit by your next higher level during the exercise.

Highly
excessive

Somewhat
too much

About the
right amount

1 2 3 4 5 6 7

_____ 14. Rate the freedom to make decisions permitted to your unit by your next higher level during the exercise.

No freedom
at all

Some but -
not enough
to do my job

About right
amount

1 2 3 4 5 6 7

_____ 15. Rate the caliber of supervision exercised over your unit by your next higher level during the exercise.

Poor

Adequate

Excellent

1 2 3 4 5 6 7

ARI Document Control
No. 77:5162C

Appendix M. Scoring of Assistant G3 Plans Solution

Scoring of Assistant G-3 Plans Solution Using CGSC Schoolhouse Solutions as Criteria
(From Krumm, Robins, and Ryan, 1973)

SUBTEST 1 FORM OF DEFENSE

Avenues of Approach (1:50,000 map)

Key Terrain Feature (1:50,000 map)

		Score	
		Correctly Drawn	Incorrectly Drawn
1. Avenue A	1. Hill mass	1	0
2. Avenue B	715, 795, 726	1	0
3. Avenue C	2. Hill 651	1	0
4. Avenue D	3. Hill 614	1	0
5. Avenue E	4. Hill 660	1	0

Main Threat (Interview record)

Form of Defense

1. Avenue A indicated as main threat approach.	1	1. Area	1
2. Avenue A not indicated as main threat approach.	0	2. Mobile	0

SUBTEST SCORE POSSIBLE 11

SUBTEST II COURSE OF ACTION

<u>Location of GOPL (1:50,000 map)</u>	<u>Score</u>	<u>Combat Power (Data Collection form)</u>	<u>Score</u>
1. Weisse Elster River	1	1. GOP	
2. 15 Km forward of FERA	1	. 2 battalions	1
3. 10-12 Km forward of FERA	0	. other Configurations	0
4. All other locations	0		
5. Omitted	0	2. FDA	
		. 4 battalions in the north	1
		. 4 battalions in the south	1
		. other Configurations	0
		3. Reserve	
		. 3 battalions	1
		. provides 2 battalions	
		GOP task force	1
		. other Configurations	0
<u>SUBTEST SCORE POSSIBLE</u>			9

Degree of Resistance (Data Collection forms)

1. 1st Bde <u>defends</u> in north	
2nd Bde <u>defends</u> in south	1
2. 1st Bde <u>delays</u> in north	
2nd Bde <u>delays</u> in south	0
3. 1st Bde <u>screens</u> in north	
1st Bde <u>screens</u> in south	0
4. GOP forces <u>delay</u>	1
5. GOP forces <u>screen</u>	0
6. GOP forces <u>defend</u>	0

SUBTEST III TASK ORGANIZATION - GRAPHIC PORTION OF DEFENSE PLAN

1. <u>COPL (1:50,000 map)</u>	Score	4. <u>Location of Reserve (1:50,000 map)</u>	Score
. 2500 meters forward FEBA or on Hills 527, 553, 543 or 547 (± 500 meters allowed).	1	. Located in accord with school solution	1
. All other locations	0	. Other locations	0
2. <u>Brigade Boundaries (1:50,000 map)</u>		5. <u>Visualized Allowable Enemy Penetrations.</u>	
a. Lateral Boundary		. Northern portion of division area -- in accord with school solution or 1 km beyond	1
. As drawn on CGSC overlay (± 1 km either side of school solution allowed)	1	. Southern portion of division area -- in accord with school solution or 1 km beyond	1
b. Rear boundary		. Other visualized penetration	0
. As drawn on CGSC overlay (± 1 km either side of school solution allowed)	1		
. Other locations	0	6. <u>Blocking Positions</u>	
3. <u>FDA Battalion Positions (1:50,000 map)</u>		. As indicated in school solution	1
a. Northern battalions		. Other positions	0
. 3 battalions at the front, 1 battalion reserve	1	7. <u>Artillery Positions</u>	
. Other configurations	0	<u>GOP (1:50,000 map)</u>	
b. Southern battalions		. Artillery Group depicted forward of FEBA	1
. 3 battalions at the front, 1 battalion reserve	1		
. Other configurations	0		

SUBTEST III (Continued)

8. Task Organization Score

Task Organization I	Task Organization II	Task Organization III	Task Organization IV
<u>GOP</u> . N - tk hvy bn, S - inf hvy bn . Other configura- tions	<u>GOP</u> . N - tk hvy bn, S - inf hvy bn . Other configura- tions	<u>GOP</u> . N - inf hvy bn, S - ACS . Other configura- tions	<u>GOP</u> . N - ACS, S - inf hvy bn . Other configura- tions
1	1	1	1
0	0	0	0
<u>1st Bde</u> . 3 inf bns, 1 tk hvy bn . Other configura- tions . Min 2 tk Co. . Other	<u>1st Bde</u> . 3 inf bns, 1 tk hvy bn . Other configura- tions . Min 2 tk Co. . Other	<u>1st Bde</u> . 3 (-) inf bns . Other configura- tions . 1 tk hvy bn . Other configura- tions . Min 2 tk Co. . Other	<u>1st Bde</u> . 3 (-) inf bns, 1 tk hvy bn . Other configura- tions . Min 2 tk Co. . Other
1	1	1	1
0	0	0	0
1	1	1	1
0	0	0	0
<u>2d Bde</u> . 3 inf bns, 1 tk hvy bn . Other configura- tions . Min of 1 tk Co. . Other	<u>2d Bde</u> . 3 (-) inf bns, 1 inf hvy bn . Other configura- tions . Min of 1 tk Co. . Other	<u>2d Bde</u> . 3 (-) inf bns, 1 inf hvy bn . Other configura- tions . Min of 1 tk Co. . Other	<u>2d Bde</u> . 3 (-) inf bns, 1 inf hvy bn . Other configura- tions . Min of 1 tk Co. . Other
1	1	1	1
0	0	0	0
1	1	1	1
0	0	0	0
<u>3d Bde</u> . 3 bn res . Other configura- tions . Tk hvy res . Other Select this TO	<u>3d Bde</u> . 3 bns . Other configura- tions . Tk hvy res . Other Select this TO	<u>3d Bde</u> . 3 bns . Other configura- tions . Tk hvy res . Other Select this TO	<u>3d Bde</u> . 3 bns . Other configura- tions . Tk hvy res . Other Select this TO
1	1	1	1
0	0	0	0
1	1	1	1
0	0	0	0
3	3	2	1

SUBTEST III (Continued)

9. Artillery Task Organization	Score	11. Engineering Task Organization	Score
GOP (data collection forms) . 1-47, 2-631 initially to GOP . 1-47, 2-631 not to GOP . 1-48, initially to GOP . 1-48, initially not to GOP . A/1-439 ADA initially to GOP . A/1-439 ADA not to GOP			
First Brigade . 1-45th assigned . 1/45th not assigned . B/1-439 ADA assigned . B/1-439 ADA not assigned	1 0 1 0	GOP . Support indicated . Support not indicated	1 0
Second Brigade . 1-46th assigned . 1-46th not assigned . A/1-439 after GOP . A/1-439 not assigned after GOP	1 0 1 0	Second Brigade . Support indicated . Support not indicated	1 0
Third Brigade . 1-47, 2-631 after GOP . 1-47, 2-631 not assigned after GOP	1 0	Third Brigade . Support indicated . Support not indicated	1 0
Division Artillery . 1-49 assigned . 1-49 not assigned . 1-48 after GOP . 1-48 not assigned after GOP	1 0 1 0	Summary Subtest I Subtest II Subtest III TOTAL POSSIBLE TEST SCORE	11 9 50 70

NAVY PERSONNEL RESEARCH AND DEVELOPMENT CENTER
SAN DIEGO, CALIFORNIA 92152-6800

26 March 1987

CAPT P. C. Nelson
Commanding Officer
Tactical Training Group, Atlantic
Virginia Beach, VA 23461-5596

Dear Captain Nelson:

The purpose of this letter is to follow up on our phone conversation yesterday regarding the Headquarters Effectiveness Assessment Tool (HEAT). The enclosed document provides our review of the materials you received from Defense Systems, Inc. (DSI). It's clear from our conversation that you're aware of many of the important issues that need to be addressed in conducting a HEAT evaluation. If no special active duty time can be obtained for your reserve detachment, conducting a single BFIT evaluation will be, in itself, a challenge. As we discussed, evaluating all three BFITs scheduled each year will require much more time than these 15 officers will have to contribute to the effort.

I hope our review adequately addresses your concerns and is helpful to you in gearing up for and conducting the upcoming BFIT. If you have any questions about our comments or if we can be of any additional assistance, please feel free to contact me (at AV 933-2396) or Dr. Susan Hearold (at AV 933-6935).



Deborah A. Mohr
Research Psychologist

Encl:

- (1) Evaluation of the Headquarters Effectiveness Assessment Tool: Defense Systems, Inc.'s Battle Force Inport Training Command and Control Evaluation Guide

**EVALUATION OF THE HEADQUARTERS EFFECTIVENESS ASSESSMENT TOOL:
DEFENSE SYSTEMS, INC.'S
BATTLE FORCE INPORT TRAINING
COMMAND AND CONTROL EVALUATION GUIDE**

The purpose of this paper is to review the deliverable provided by Defense Systems, Inc. (DSI) to Tactical Training Group, Atlantic (TACTRAGRULANT) that describes the application of the Headquarters Effectiveness Assessment Tool (HEAT) to evaluate performance during Commander, Second Fleet (COMSECONDFLT) Battle Force Inport Training (BFIT) exercises. This review focuses on two factors: (1) the HEAT methodology itself, and (2) the feasibility of applying the methodology to BFITs using DSI's guide.

HEAT as an Assessment Methodology

The issues discussed here focus on the advantages and disadvantages of HEAT as a methodology for battle force/battle group (BF/BG) performance assessment.

Advantages

The advantages of the HEAT methodology are: (1) its systematic model-based nature, (2) its comprehensive assessment of planning functions, (3) the quantitative measures it provides, (4) its use of multiple incidents on which to base performance scores, and (5) the lack of alternative assessment methodologies. These are discussed below.

1. Systematic model-based methodology. HEAT is based on a model of headquarters organizations as adaptive control systems in which control is exercised through planning and execution cycles. HEAT measures follow directly from this model and are designed to tap the phases of the adaptive control process. These phases are: monitor, understand, develop alternative actions, predict, decide, and direct. Overall measures of effectiveness (MOEs) tap the length of time plans remain valid and supporting measures assess command performance in the adaptive control phases. As a model-based methodology, HEAT's measures are systematic in nature and are designed to assess performance in each of the component phases of the model. This contrasts with methodologies that are not model based and that tend to develop measures in a haphazard or random manner.

2. Comprehensive coverage of planning functions. HEAT provides two primary measures of effectiveness that focus on the viability of plans and many supporting process measures to be used for diagnosis. The various phases of planning are well covered by HEAT's many measures. Understanding the current situation and expectations about the future situation, generating options to respond to expected futures, predicting the impact of potential actions, selecting the appropriate option, and directing execution of the selected option are all assessed.

3. Quantitative measures. HEAT provides quantitative (numerical) scores for various headquarters commands on each of the measures. HEAT measures are either percentage scores (e.g., percent of directions queried by recipients) or time scores (e.g., average time to issue a directive after a decision has been made). Both allow aggregation across warfare areas or scenario phases or to higher level headquarters commands and permit comparison of performance among participating commands. Some (but not all) of these quantitative measures eliminate the need to rely on observers' subjective assessments of the battle group's performance in the command and control arena.

4. Measures based on multiple incidents. Calculation of many of the HEAT measures is based on more than one incident during the exercise. For example, understanding of the enemy situation is based, not on one time period (e.g., at problem start), but rather on every time the understanding changes (due either to additional information or a situation change). Thus, the resulting scores reflect performance on a broad sample of understandings and are therefore more likely to reflect true performance in each assessment area.

5. Lack of alternative methodologies. There are few other methodologies available for assessing BF/BG performance, particularly in the cross-warfare coordination and command and control areas. Most existing MOEs and MOPs assess weapons and targeting effectiveness. Such measures focus on the end result (i.e., PK or hit or miss) of the complex command and control efforts required in effective naval warfare. Limited efforts have been made toward developing subjective assessment procedures for BF/BG command and control (e.g., BGE or ORA assessment methodologies). HEAT attempts to go beyond these measures to provide quantitative assessments of the process of naval warfare.

Disadvantages

The disadvantages of the HEAT methodology are: (1) its focus on process at the exclusion of outcomes, (2) the lack of assessment of platform performance and execution of tactical decisions, (3) the subjectivity of measures and performance standards, (4) the quantity of data collection and analysis required, and (5) its unknown psychometric qualities.

1. Process orientation rather than outcome orientation. As mentioned, HEAT focuses on the process of naval warfare, specifically the planning and execution cycles of an adaptive control system. No attempt is made to assess the quality of tactical decisions, the outcomes of tactical decisions, or the extent to which executed plans are effective. Theoretically, the "correct" plan, given available information about the tactical situation, might be selected and implemented, but fail to achieve expected results. A HEAT assessment would not give high scores for the innovation, intuition, and creativity that are sometimes critical to successful warfare.

In addition, HEAT's primary MOE assesses the length of time plans remain in effect and is based on the premise that more effective planning results in plans that are viable for longer periods of time. However, headquarters organizations may stay with an implemented plan beyond the time it is appropriate (for example, due to lack of knowledge about the true tactical situation or inability to adapt to changes in the environment or situation). HEAT measures would not reflect such situations when rigidity limits the BF/BG's ability to adapt to the changing tactical situation. The command organization would receive a positive score on the MOE despite the fact that their true performance was less than optimal.

2. Platform performance and execution of tactical decisions not fully assessed. HEAT provides no measures of the performance of many BG participants. The numerous air, surface, and subsurface units that support the command levels being assessed are not themselves evaluated. Likewise, the performance of warfare area commanders and supporting element coordinators is not assessed. There are two disadvantages resulting from this. First, if command level (BF/BG) problems are identified, no subsequent analysis can be conducted to determine, for example, the platform that provided incorrect or late information or the warfare area commander who failed to interpret a directive correctly. Second, all BG participants, from CIC watch standers to the BG commander, spend

considerable time preparing for and taking part in BFIT exercises. To expect this level of support without providing any systematic feedback on subordinate performance, limits the potential gain from participation in such exercises. Providing performance feedback to all participants would improve the learning experience gained from BFIT.

As mentioned earlier, HEAT does not assess the execution of selected plans. Feedback on the outcome of various planning cycles would provide additional information to participants (from the platform level to the BF/BG command level) on their performance during the exercise.

3. Subjectivity of measures and performance standards. While all HEAT measures provide quantitative scores, many are, nonetheless, based on subjective judgment. The experience, training, and attentiveness of observers collecting required data may all influence resulting HEAT scores. It is critical that observers/data collectors have knowledge and/or experience in battle group organization, procedures and tactics; understand HEAT rationale and procedures; and remain attentive during the complex and busy exercise execution. In addition, since some observers may be selected from participating commands, they must maintain objectivity in their observations as much as possible. The need for consistency in HEAT procedures is critical given the changing group of observers and analysts assessing each BFIT.

Several of the HEAT measures require performance standards to be set against which observed performance is compared (MOEs B-2, B-4, B-19, and B-20). These standards are typically set by HEAT users (e.g., COMSECONDFLT or BF/BG CWCs) and may or may not be accepted or acceptable standards. They require the subjective judgment of the user and must be re-evaluated prior to each HEAT application. Because no objective standards exist in most areas, they could be set in a way that biases assessment results (for example, by setting "easy" standards to make the battle group look good). While performance standards should be adapted to the unique needs and capabilities of each BF/BG, the ability to compare resulting HEAT scores from one BFIT to the next is then compromised.

Certain HEAT measures are based on DSI specified standards. For example, higher scores are given when more than one planner participates in options generation, but DSI admits that "any predictable relation between this number and overall effectiveness remains to be determined (p. II-9)". The individuals comprising the planning team undoubtedly have more impact on effectiveness than the mere number of planners. One effective planner might win the war, whereas a greater number of untrained, unskilled, or uninterested planners might hand victory to the enemy. Care should be taken in interpreting results when such standards are not based on objective performance requirements nor on proven relationships with battle group effectiveness.

Related to the subjectivity of standards is another issue with perhaps more serious implications. When individuals or groups are aware of the measures being used to evaluate their performance, they may focus their efforts on these measured factors in order to receive a good evaluation. This is not a problem if measures cover all aspects of performance and are well constructed. However, if key performance areas are not assessed or if measures do not really tap the desired performance, measured performance may be maximized to the detriment of overall true performance. For example, MOE P-8 provides higher scores when more than one option is considered in developing plans. BF/BG staffs might be tempted to consider unrealistic, over-simplified, or meaningless options simply to satisfy what becomes a performance requirement or goal once measured. Since the MOE does not consider the quality or realism of the options considered, an inappropriate high score might be given in

such situations. Efforts should be made to ensure that command staffs being evaluated do not attempt to subvert or undermine the assessment program.

4. Quantity of data collection and analysis efforts required. With 20 planning (CPX) phase measures and 29 battle problem measures, a considerable data collection effort is required. The mere number of measures greatly underestimates this effort, though. Many of these measures actually consist of separate scores for a number of data categories (e.g., 7 functional warfare areas for MOE B-1, 6 characteristics of own forces for MOE B-3, and 4 aspects of the enemy situation for MOE P-17). If data are collected in all data categories shown in Tables III-1 and III-2 (though most likely not all categories will be used in every BFIT), a total of 153 planning measures and 210 battle problem measures will result. In addition, scores are often based on multiple planning or battle problem incidents. Thus, any one measure may require a wealth of data collection. For example, monitoring own force accuracy (MOE B-3) is based on a command center's estimates of 6 characteristics of each own force unit and is scored either whenever any of those estimates changes or at periodic predetermined times during the exercise.

Analysis of collected data and interpretation of results are, likewise, large and complex tasks. Following creation of an overall timeline of significant battle problem events and control cycles, data must be transferred from observers' records to data sheets, combined within like data categories (e.g., warfare area or exercise event), analyzed and transformed to HEAT scores, and compiled into meaningful feedback. The large number of individual measures that will result from a full scale HEAT assessment must be well integrated and interpreted to be useful to participating staffs and to COMSECONDFLT. In order to provide timely, useful performance feedback to participating battle staffs, considerable effort will need to be devoted to data analysis and interpretation soon after FINEX.

5. Unknown psychometric properties. There are two primary psychometric qualities that are important in evaluating tests and other assessment techniques. Reliability refers to the extent to which similar scores would result if the same entity were repeatedly evaluated. In the present context, a reliable assessment technique would provide consistent judgments of BF/BG performance. Unreliable tests do not provide consistent information about the entity being evaluated and therefore are not useful. Validity refers to the extent to which the test or technique measures what it intends to measure. In looking at headquarters assessment methods, valid methods would actually measure headquarters effectiveness. Invalid measures might reflect BF/BG composition or exercise scenario more than true effectiveness.

No evidence beyond information of its previous applications has been provided by DSI to support the reliability or validity of the HEAT methodology. Looking first at reliability, HEAT scores may vary from administration to administration due to: the subjectivity of measures, changes in battle group composition, assignment of different observers to battle staffs, changes in exercise scenario, and modification of performance standards. Any or all of these factors may influence HEAT scores, thereby limiting their overall reliability and the ability to compare scores.

Because no validity information is provided in the guide for the HEAT methodology, no conclusions can be drawn on the true validity of the measures. The extent to which HEAT measures assess headquarters' performance will depend on the extent to which the exercise itself is typical of actual or potential battle group scenarios and the extent to which the battle group composition, organization, and procedures reflect real world BF/BG operations.

Without additional psychometric data about the HEAT methodology, interpretation of differences among scores and differences between BFITs is ambiguous. For example, it is unclear at this time whether a difference of .1 or .9 on standardized scores is required to conclude that meaningful differences in true performance exist. While some statistical tests are suggested for determining the statistical significance of differences, the meaningfulness of such differences cannot be determined. If two commands receive HEAT scores that differ by .1, analysts and users can say little about the meaning of that difference. In addition, until many similar battle groups have participated in similar BFIT exercises and been evaluated with the HEAT methodology, no baseline of typical performance will be available against which current performance can be compared. This lack of normative data limits the current usefulness of the methodology, but will be mitigated with additional applications of HEAT to BFIT exercises and the development of a HEAT data base compiling results.

Finally, given these reliability and validity questions, the usefulness of HEAT results must be examined. Assessment results must meet the needs of the user. Presumably BFITs are designed as training exercises to assess BF/BG capabilities and to identify deficiencies. If HEAT results cannot identify performance problems or do not reflect the BF/BG's true performance capabilities (i.e., in real world situations), little will be gained from the effort.

Feasibility of Applying the HEAT Methodology to BFITs

The issues discussed here focus on the extent to which TACTRAGRULANT's reserve detachment for HEAT can effectively apply the methodology to BFITs using DSI's deliverable. Our comments are addressed to two major issues: the completeness and clarity of the package and the time requirements to implement the methodology.

DSI's HEAT Evaluation Guide

A thorough review of the Battle Force Inport Training Command and Control Evaluation Guide prepared by DSI reveals more comprehensive, clear coverage of the application of HEAT to BFITs than in previous DSI reports or manuals. The rationale and theory underlying HEAT are adequately covered, as is the adaptive control cycle. The importance of early and thorough planning prior to a BFIT evolution are properly stressed and specific tasks to be accomplished before the assessment begins are specified. The guidelines for observers (pp. VI-2 to VI-17) may appear simplistic and common-sense, but should not be neglected.

The sections on data collection and analysis have been revised and are more straightforward and comprehensive. Data collection sheets are more direct and less ambiguous and will likely provide more accurate, useful information than previous versions of these sheets. The data reduction process has been simplified and is described in the guide in fairly explicit detail. Transformation of raw scores to standard scores, a complex procedure described in previous HEAT manuals that adds little to the assessment effort, has been eliminated. Competent data analysts will probably be able to go through this process effectively using the guide. The guide has two specific deficiencies in this area, however. First, no procedures are provided for conducting or interpreting the statistical tests that are suggested. Unless DSI incorporates such procedures in the guide, observers/analysts will need to refer to one of the statistics books referenced in the text for assistance in carrying out these tests. Second, no data categories are specified in Table III-2 for MOEs dealing with the enemy situation (MOEs B-2 and B-21). The Enemy Location Data Sheet (Figure

VII-6) implies that enemy location may be the only data category for these MOEs. These deficiencies should be corrected prior to conducting a HEAT assessment. Both data collection and analysis will likely be slow and time-consuming in initial applications with the need for continual reliance on the evaluation guide. As analysts gain experience, these time requirements will drop.

The quality of the assessment and resulting feedback will depend to a large extent on the experience and competence of the individuals in the reserve detachment. Ideally, these officers should have operational battle group experience and good quantitative and writing skills. The interpretation of HEAT scores and the presentation of observers' insights will depend on the skills and experience these officers bring to the task.

Time Requirements

As mentioned previously, data collection and analysis efforts required to successfully conduct a HEAT evaluation are substantial. With unlimited resources (time and personnel), HEAT may provide valuable and meaningful results. Your letter indicates, however, that the reserve detachment of 15 officers will have to do all HEAT planning, data collection, and analysis functions. It appears that DSI is not fully aware of this since they mention the use of battle staff observers (during the planning phase) and data analysts (to assist in computation of HEAT scores) in addition to the reserve officers.

The effective application of the HEAT methodology to Second Fleet BFTTs will depend on adequate support. This includes support from COMSECONDFLT, from you and your staff, from the reserve officers conducting the assessment, as well as from battle force staffs being evaluated. The importance of support cannot be underestimated. If battle staffs being observed attempt to avoid or thwart the assessment effort, results will be inaccurate, incomplete, or meaningless. COMSECONDFLT's support will go far toward ensuring that BF/BG staffs also support the assessment. COMSECONDFLT may also be able to ensure that battle staff personnel are provided to serve as participant observers during the planning phase and that other personnel are available to assist in the data analysis effort.

With a reserve detachment comprised of 15 competent officers, an adequate HEAT evaluation can probably be conducted. The personnel requirements and logistics of the evaluation are substantial given three or more battle groups in three geographically separated locations. Such an evaluation would be easier and feedback more timely with additional support. Prior to the exercise planning phase, time is required for all observers and analysts to be trained and to become thoroughly familiar with the HEAT methodology and procedures. In addition to reviewing the evaluation guide in detail, it is suggested that all individuals involved in the assessment practice data collection, reduction, and analysis to the greatest extent possible. Part of this practice should include a consistency check to determine the extent to which observers complete data sheets and calculate HEAT scores in the same way. HEAT results will be unreliable and meaningless if observers do not follow standardized procedures. Through early training and practice, problems that may surface can be addressed before the exercise is underway and while adequate time is available to resolve them.

During the planning phase, one observer is required for each battle group command staff and for the battle force command staff, at a minimum. DSI estimates that 3 to 4 hours per week per command staff are required during the planning phase. However, it may be difficult to reconstruct important details of the planning process in several hours if the BF/BG staff has spent substantial time in planning since a previous visit. A preferable

means of accommodating the situation would be to assign responsibility for planning phase data collection to a team comprised of an individual from the command staff and one or more reserve officers who make periodic visits to the command staffs. In this way, the participant observer assigned from the battle force/group staff could collect needed data and provide insights to planning processes that occur while the reservists are not present. COMSECONDFLT's support is required to undertake such an effort and battle staff participant observers need to attend HEAT training along with other observers. Because the amount of data collection required during the planning phase is significant, these observers will need to have adequate time allocated for this effort.

During the battle problem, observers need to monitor the BF/BG command staffs throughout the exercise and to spend additional time recording observations and tabulating data. Given the size of the data collection requirement, more than one observer should be assigned to each of the BF/BG staffs. Ideally, there should be some continuity of observers during the two phases. Observers should be assigned to the same staffs during the planning and battle problem phases. This continuity would improve the observers' ability to develop meaningful insights and would enhance command staff acceptance of the observers' purpose and subsequent insights. Given the limited active duty time of the assigned reserve officers, data collection may have to be reduced to fit the available resources during each phase.

Additional reservists are required to collect ground truth data against which battle staff perceptions and evaluations will be compared. At least one officer should be assigned to observe COMORANGE for this purpose. Efforts could be made to enlist the assistance of Fleet Combat Training Center, Atlantic (FCTCLANT) personnel to provide some record of the ground truth positions of own force and enemy units throughout the exercise. The timing of certain ground truth data collection must correspond to collection of parallel data on perceptions of the situation. For example, perceptions of the location of enemy units must be compared to actual enemy unit locations at similar times.

After the planning phase and particularly after FINEX, observers and analysts face a significant task. The collected data must be tabulated, analyzed, and brought together to provide meaningful feedback. The value of performance feedback decreases as the time delay after exercise participation increases. Thus, it is important to process and summarize HEAT data and to provide feedback to participants in a timely manner. A post-exercise brief summarizing important analysis results and insights should be given as soon after FINEX as feasible, followed shortly thereafter with a more complete written report of the assessment effort. Users' expectations about the products of the HEAT assessment should be kept realistic. Comments provided by the evaluation guide on the form and content of the briefing and report are straightforward, yet preparing these products may require substantial resources. In order to remain within the capabilities of the available personnel, some limitation in the scope of these products may be needed. The reserve officers preparing these products should be provided clerical support during this phase.

A long term solution to the problems of quick turn-around on performance feedback may be to develop an automated HEAT report package. Such a package would provide a general pre-written HEAT assessment report that consists of a standard format into which results from the current BFIT can be easily inserted. With some tailoring of the document to fit each exercise and insertion of results and insights, a feedback report could be generated.

In order to accomplish a successful HEAT evaluation, considerable efforts will be required to manage the program and to provide continuity from one BFIT to the next. One officer, either reserve or active duty, should be assigned responsibility for coordinating all of

the activities required by this methodology and resolving problems that arise. Without sufficient leadership and direction, the assessment effort may fail. The same officer can provide continuity between exercises. Each HEAT application will be as difficult as the first unless some means are developed to maintain and build on knowledge and experience gained from previous exercises.

Suggestions

This section summarizes suggestions made in the above discussion and presents additional ideas to consider in applying HEAT to upcoming BFITs.

1. Ensure that top down support (i.e., originating with COMSECONDFLT) exists for using the HEAT methodology and that this support is communicated to BF/BG command staffs, HEAT observers (both reserve and BF/BG staff observers), and data analysts. This support should include the provision of resources required to make the assessment complete and meaningful.

2. Ensure that observers and analysts assigned to the HEAT assessment have the experience and skills required to do the best job possible. Ideally, continuity should be maximized in selecting and assigning individuals to commands during the two exercise phases. In addition, all of these individuals should be well trained in the methodology and should participate in practice drills prior to the BFIT. Given that the composition of the reserve detachment will change from one BFIT to the next, stable direction and leadership should be provided to ensure that all new observers are brought up to speed and that the application of the methodology is consistent, thorough, and correct.

3. Be aware that, although many HEAT scores are numerical, they are nonetheless subjective in nature. Assigning qualified officers to this effort and training them in the proper use of the HEAT methodology will improve the meaningfulness of such subjective assessments. Where performance standards must be set, keep in mind the impact of these selected standards on resulting HEAT scores.

4. Consider limiting the scope of the assessment effort to keep it within the capabilities of available observers and analysts. This might include selecting a subset of measures (and/or a subset of the data categories shown in Tables III-1 and III-2 for selected measures) that focus most directly on the objectives of the exercise and yet still meet COMSECONDFLT's evaluation needs. It might mean assessing the adaptive control cycle associated with selected key developments in the scenario rather than assessing all control cycles.

5. Use the TACTRAGRULANT reserve detachment's initial attempt to perform the HEAT assessment as a trial evaluation. Monitor the time and resources required to prepare for and conduct the assessment and weigh these against the feedback and products provided by the assessment effort. Keep in mind that with additional experience using the HEAT methodology, the time requirements may decrease, while the quality of results may increase. The initial assessment may reveal areas in which cost effective changes can be made or ways in which the process or product can be improved. Alternately, it may show that the results are not worth the costs required to conduct the assessment. Such an evaluation will have to be subjective, based on the goals for BFIT performance assessment and the availability of necessary resources.

6. Continue supplementing the HEAT methodology with additional efforts that more directly assess tactical proficiency (e.g., observers' subjective assessments). In the long run, other assessment techniques should also be considered to provide performance feedback to participants throughout the BF/BG. Some of the HEAT measures used for BF/BG assessments, can be applied to other levels as well. For example, monitoring accuracy/timeliness (MOE B-2) is a measure applicable to all platforms participating in the exercise. While such efforts are beyond the capabilities of the current BFIT evaluation resources, they should be considered as potential means of improving the effectiveness of the entire battle force or group.