U.S. Army Research Institute
for the Behavioral and Social Sciences

Research Report 1764

Refinement of Prototype Staff Evaluation Methods for Future Forces: A Focus on Automated Measures

May H. Throne and William T. Holden, Jr.
Human Resources Research Organization

Carl W. Lickteig
U.S. Army Research Institute

November 2000

Approved for public release; distribution is unlimited.

20001120 099

U.S. Army Research Institute for the Behavioral and Social Sciences

A Directorate of the U.S. Total Army Personnel Command

EDGAR M. JOHNSON
Director

Research accomplished under contract
for the Department of the Army

Human Resources Research Organization

Technical review by

Larry L. Meliza
David Bessemer

NOTICES

DISTRIBUTION: Primary distribution of this Research 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: TAPC-ARI-PO, 5001 Eisenhower Ave., Alexandria, VA 22333-5600.

FINAL DISPOSITION: This Research 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 Research Report are not to be construed as an official Department of the Army position, unless so designated by other authorized documents.

REPORT DOCUMENTATION PAGE

1. REPORT DATE (dd-mm-yy) November 2000	2. REPORT TYPE Final	3. DATES COVERED (from. . . to) July 1999 - August 2000
4. TITLE AND SUBTITLE Refinement of Prototype Staff Evaluation Methods for Future Forces: A Focus on Automated Measures		5a. CONTRACT OR GRANT NUMBER DASW01-99-D-0012 DO #2 5b. PROGRAM ELEMENT NUMBER 0602785A
6. AUTHOR(S) May H. Throne and William T. Holden, Jr. (HumRRO); Carl W. Lickteig (U.S. Army Research Institute)		5c. PROJECT NUMBER A790 5d. TASK NUMBER 2228 5e. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Human Resources Research Organization (HumRRO) 66 Canal Center Plaza, Suite 400 Alexandria, VA 22314 U.S. Army Research Institute for the Behavioral and Social Sciences, Attn: TAPC-ARI-IK 2423 Morande Street, Fort Knox, KY 40121		8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Research Institute for the Behavioral and Social Sciences Attn: TAPC-ARI-IK 5001 Eisenhower Avenue Alexandria, VA 22333-5600		10. MONITOR ACRONYM ARI 11. MONITOR REPORT NUMBER Research Report 1764

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

Carl W. Lickteig, Contracting Officer's Representative

14. ABSTRACT *(Maximum 200 words)*:

This report examines the use of digital information systems and automated measures of human performance to improve staff training and performance assessment. This work began with a review of research literature and technical documentation related to team performance and assessment, operations in digital environments, and automated performance data collection. A general design for staff performance assessment was formulated, based on findings of the literature review. An opportunity to implement this design was provided by an Army Concept Experimentation Program (CEP), the Battle Command Reengineering (BCR) IV, which took place in April 2000. By participating in the BCR IV, researchers had the opportunity to conduct a trial implementation of the automated measures of performance assessment. Coordination between the U.S. Army Research Institute for the Behavioral and Social Sciences and the Mounted Maneuver Battlespace Lab (MMBL) at Fort Knox, Kentucky, enabled the two organizations to work together as a team to accomplish multiple goals. This report describes the development of prototype automated measures, the results of their use during the BCR IV, and lessons learned for future staff performance assessment efforts.

15. SUBJECT TERMS

Battle Command Reengineering; Staff Processes; Team Performance; Mounted Maneuver Battlespace Laboratory; Automated Measurement; C^4I

SECURITY CLASSIFICATION OF			19. LIMITATION OF ABSTRACT	20. NUMBER OF PAGES	21. RESPONSIBLE PERSON (Name and Telephone Number)
16. REPORT Unclassified	17. ABSTRACT Unclassified	18. THIS PAGE Unclassified	Unlimited	107	Dr. Carl W. Lickteig DSN 464-2613

NSN 7540-01-280-5500

Standard Form 298

Research Report 1764

Refinement of Prototype Staff Evaluation Methods for Future Forces: A Focus on Automated Measures

May H. Throne and William T. Holden, Jr.
Human Resources Research Organization

Carl W. Lickteig
U.S. Army Research Institute

Armored Forces Research Unit
Barbara A. Black, Chief

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

November 2000

Army Project Number
2O262785A790

Personnel Performance and Training Technology

Approved for public release; distribution is unlimited.

FOREWORD

The Future Battlefield Conditions (FBC) Team of the Armored Forces Research Unit, U.S. Army Research Institute for the Behavioral and Social Sciences (ARI) has a Science and Technology Objective (STO) entitled "Force XXI Training Strategies." This STO is also reflected in the FBC work package (2228) FASTRAIN: Force XXI Training Methods and Strategies. Recent work under this work package has involved research and development concerning training for digital staffs. In order to continue this work, a contract entitled "Refinement of Methods for the Training and Assessment of Digital Staffs" was issued. The major purpose of this effort was to refine the prototype training and assessment techniques for future information age staffs that had been developed during a prior research effort.

This report concerns assessment for future battle staffs equipped with digital command, control, communications, computers, and intelligence (C^4I) systems. It describes an examination of theories of team processes and assessment techniques applicable to information age staffs. The report documents the design and development of prototype automated assessment methods based on those theories. The report examines implementation of the prototype evaluation techniques in a battalion level Battle Command Reengineering (BCR) IV experiment, conducted by the Mounted Maneuver Battlespace Lab (MMBL) located at Fort Knox, Kentucky. Lessons learned are documented as a guide for further development of automated measures of staff performance.

At least two major audiences may be interested in this report. Researchers interested in the area of training evaluation for information age staffs will find an examination of a set of prototype automated measures of staff performance that were developed during this research. Also, the report may be of interest to digital C^4I system developers, in that it describes an effort to implement automated measures of staff performance based on the data generated and stored by digital systems. Thus, this report may prove useful in future research and development efforts for assessing information age staff performance and in designing automated performance feedback mechanisms for staffs.

The prototype products developed under this effort are documented in a six-volume set of materials entitled *Training and Measurement Support Package, Battle Command Reengineering IV, Mounted Maneuver Battlespace Lab* (ARI, 2000) available from the MMBL. Evaluation findings from this effort are included in the MMBL's *Battle Lab Experiment Final Report (BLEFR) for Battle Command Reengineering, Phase IV* (in preparation).

The research reflected in this report was briefed to MMBL sponsors in a final In Progress Review, held at Armored Forces Research Unit, Fort Knox, Kentucky, on 13 July 2000.

Zita M. Simutis

ZITA M. SIMUTIS
Technical Director

ACKNOWLEDGMENTS

This report represents the efforts of an integrated team of research scientists, military experts, training developers, simulation technology experts, and administrative support personnel. A highly skilled and dedicated group of professionals supported the authors on this effort.

The team members directly supporting the U.S. Army Research Institute for the Behavioral and Social Sciences included: military subject matter expertise from Mr. Neff Jenkins (Litton PRC), and Mr. Mike Cobb (Human Resources Research Organization [HumRRO]); measurement expertise from Dr. Laura Ford (HumRRO); simulations expertise from Mr. Bud Dannemiller (Litton PRC); and instructional technology support from Ms. Kathy Horn (Litton PRC). Ms. Charlotte Campbell (HumRRO) was the Program Manager for this project. Her contributions went far beyond project supervision. She was involved in all project decisions, contributed her expertise as a Research Psychologist, and was a reviewer of this report.

Additionally, we had support and guidance from a variety of individuals and government organizations, including:

- Mounted Maneuver Battlespace Lab, Fort Knox, Kentucky

  Colonel Richard Savage, Chief
  Major Joe Burns, Chief, Advanced Concepts Research
  Captain Dan Ray, Concept Integration Officer

- Illinois Institute of Technology Research Institute (IITRI)/AB Technologies

  Mr. Pat Ritter, Program Director for Research
  Mr. Jim Lewis, Senior Systems Analyst

- Lockheed Martin-Marietta

  Mr. Don Appler, Site Manager
  Mr. Dan Schultz, Battle Master
  Mr. Charles West, Assistant Battle Master
  Mr. Paul Monday, Analyst
  Mr. Paul Colonna, Analyst

- Officers, Non-Commissioned Officers, and Soldiers of 2nd Armored Cavalry Regiment, Fort Polk, Louisiana

REFINEMENT OF PROTOTYPE STAFF EVALUATION METHODS FOR FUTURE FORCES: A FOCUS ON AUTOMATED MEASURES

EXECUTIVE SUMMARY

Research Requirement:

The U.S. Army is currently developing and fielding information systems for the digital battlefield of the future. In support of this effort, the U.S. Army Research Institute for the Behavioral and Social Sciences (ARI), Armored Forces Research Unit, Future Battlefield Conditions Team is engaged in the design and development of training and performance evaluation techniques. For this project, ARI's objective was to refine strategies for automated assessment of staff performance in the future digital tactical operations center at brigade and below. In order to accomplish this objective, the Project Team designed and developed a prototype automated measures package for assessment of staff processes.

The prototype automated measures were implemented during the Battle Command Reengineering (BCR) IV Concept Experimentation Program experiment conducted by the Mounted Maneuver Battlespace Lab (MMBL) at Fort Knox, Kentucky. The purpose for ARI's participation in this experiment was to refine the prototype automated measures package developed for the MMBL's research efforts, and to support the Army's need to gain additional information on how to meet future staff evaluation requirements.

Procedure:

The Project Team reviewed available literature regarding staff processes and measures developed to assess them, especially automated measures. This literature review provided the basis for decisions concerning staff processes to measure, measures design, development, and implementation, and analysis procedures. An extensive front-end analysis was conducted to define future battalion-level staff processes to be measured. After analysis, candidate measures were designed and then developed in conjunction with support from a Subject Matter Expert Advisory Group.

The prototype automated measures package was implemented during the BCR IV experiment, which took place 3 April through 19 April 2000. The experiment was conducted in the MMBL Mounted Warfare Test Bed at Fort Knox with the 2nd Armored Cavalry Regiment from Fort Polk participating. The major research products associated with training and evaluation for the MMBL implementation are presented in the six-volume set of materials entitled *Training and Measurement Support Package, Battle Command Reengineering IV, Mounted Maneuver Battlespace Lab* (ARI, 2000).

Findings:

The Project Team made an effort to develop automated measures of staff performance that could take advantage of the analytical power and processing speed of advanced command, control, communications, computers, and intelligence (C^4I) systems to provide real or near real-time feedback to training participants. This effort was only partially successful. Considerable research remains to be done on identifying both staff processes and tasks that can be measured through automated means and developing output formats for performance feedback to future battle staffs operating advanced C^4I systems. Further research is also required to determine the specific methods and measures needed to extract data from advanced C^4I systems or simulations in an easily interpretable format suitable for performance feedback.

Utilization of Findings:

These findings provide prototype examples and lessons learned on the design and development of automated measures of performance for battalion and brigade-level staffs that are equipped with advanced digital C^4I systems. The findings also provide examples of automated measures that may be required to assess embedded training and performance for future forces, including Brigade Combat Teams. The findings were provided to the MMBL for inclusion in the final BCR IV experiment report (MMBL, in preparation). The audience who may find the information contained in this report beneficial includes readers interested in the area of training evaluation for information age staffs, and digital C^4I systems designers and developers.

REFINEMENT OF PROTOTYPE STAFF EVALUATION METHODS FOR FUTURE FORCES: A FOCUS ON AUTOMATED MEASURES

CONTENTS

Page

INTRODUCTION 1

Organization of the Report 3
Battle Command Reengineering Setting 3

Experimental Objectives 4
Surrogate Command, Control, Communications, and Computers System 5
Instrumentation and Data Collection 6
Army Battalion Staff Participation 8

Previous Findings 8

Staff Processes 10
Measures of Staff Processes 13
Automated Measures of Staff Performance 16
Summary 20

METHOD 20

Measures Development Process 20

Front-End Analysis 21
Design 24
Development 26

Measures Implementation Process 30

Participants 30
Battle Command Reengineering Missions 31
Implementation 32

BATTLE COMMAND REENGINEERING IV RESULTS AND DISCUSSION 33

Adaptability 34

Terrain Analysis 34
Summary of Adaptability 36

CONTENTS (Continued)

Page

Performance Monitoring and Feedback 36

Spot Report Use 37
Summary of Performance Monitoring and Feedback 38

Shared Situational Awareness 39

Map Area 39
Summary of Shared Situational Awareness 42

Communication 43

Whiteboard Use 43
Summary of Communication 45

Coordination 45

Fire Support Coordination 45
Opposing Forces Destruction 47
Summary of Coordination 48

Decision-Making 48

Unmanned Aerial Vehicle Effectiveness 49
Summary of Decision-Making 50

Summary 50

BATTLE COMMAND REENGINEERING IV IMPLEMENTATION LESSONS LEARNED 51

Automated Staff Performance Feedback 51
Measurement Framework 52
Balanced Measures 52
Measures Development 52
Data Reduction 54
Pictorial Output Formats 54
Automated Flags 54
Summary 55

FUTURE EFFORTS 55

Page

Measurement Framework 57
Balanced Measures 57
Command, Control, Communications, Computers, and Intelligence Instrumentation 57
Digital Integration 58
Multidisciplinary Teams 58
Iterative Processes 58
Data Flags 59

Automated System Flags 59
User-Defined Flags 59

Data Reduction 59
Output Formats 60
Performance Feedback 60
Training Audience 60
Summary 61

REFERENCES 63

APPENDIX A. LIST OF ACRONYMS A-1

B. PROTOTYPE AUTOMATED MEASUREMENT PACKAGE B-1

LIST OF TABLES

Table 1. Cannon-Bowers et al.'s Team Process Skill Dimension Definitions 11

2. Classification of Skill Dimensions by Teamwork Model 12

3. Reengineered Battle Command (RBC) III Issues and Automated Measures of Performance 18

4. Team Process Skill Dimensions and Associated Behaviors 22

5. Candidate Automated Measures 24

6. Automated Measures Developed for Battle Command Reengineering IV by Skill Dimensions 27

7. Automated Flags for the Data Collection and Analysis System 29

8. Staff Member Call Signs and Titles 31

9. Battle Command Reengineering IV Experiment Missions 34

Page

10. Frequency and Duration of the Terrain Intervisibility Tool Use for Staff Members by Mission 36
11. Frequency and Duration of Spot Report Use for Staff Members by Mission 37
12. Average Km^2 of Battlefield Displayed by Each Staff Member for Each Mission 40
13. Number of Whiteboard Files Residing on Each Staff Member's Workstation by Mission 44
14. Frequency and Percentage of Opposing Forces (OPFOR) Kills Due to Direct and Indirect Fires by Mission 46
15. Cumulative Rate of Opposing Forces (OPFOR) Losses Every 20 Minutes for Each Mission 47
16. Summary Table of Unmanned Aerial Vehicle (UAV) Effectiveness 49
17. Summary of Implementation Lessons Learned 51
18. Summary of Lessons Learned for Future Efforts 56

LIST OF FIGURES

Figure 1. Battle Command Reengineering IV staff structure 5
2. Surrogate command, control, communications, and computers system setup 6
3. Battle Command Reengineering IV data collection system 7
4. Sample command and control battlefield operating system measures 17
5. Exercise flag log for the Battle Command Reengineering IV 30
6. Battle Command Reengineering IV terrain map 32
7. Average duration of each use of the Field of View tool, by mission 35
8. Average duration of each use of the Snail tool, by mission 35
9. Average duration of each use of the Stealth tool, by mission 35
10. Average duration of each use of the Terrain Intervisibility tool, by mission 35
11. Average frequency of spot report use for commander and staff by mission 38
12. Average frequency of spot report use for each node by mission 38

Page

13. Center points for each staff member's Plan View Display at first indirect fire for Mission 2, with the squadron commander's and a company commander's views 41

14. Center points for each staff member's Plan View Display at start of exercise for Mission 2, with the squadron commander's and a company commander's views 41

15. Center points for each staff member's Plan View Display at first direct fire for Mission 2, with the squadron commander's and a company commander's views 42

16. Center points for each staff member's Plan View Display at end of exercise for Mission 2, with the squadron commander's and a company commander's views 42

17. Average number of Whiteboard files for each node by mission 44

18. Frequency of opposing forces (OPFOR) kills due to direct and indirect fire for each mission 46

19. Cumulative rate of opposing forces (OPFOR) destruction for each mission over 20 minute time periods 48

20. Unmanned aerial vehicle (UAV) flight routes and opposing forces (OPFOR) locations for Mission 2 50

REFINEMENT OF PROTOTYPE STAFF EVALUATION METHODS FOR FUTURE FORCES: A FOCUS ON AUTOMATED MEASURES

Introduction

The Army's growing reliance on computer-mediated work underscores the potential of digital technologies to automatically collect and analyze real-world performance data (Caldera & Reimer, 1999). As command, control, communications, computers, and intelligence (C^4I) systems become more integral to the performance of individual and collective tasks, the human-computer interactions associated with these systems become more critical and collectible. The focus of this report, therefore, is on the use of digital information systems and automated measures of human performance to improve staff training and performance assessment.

With the increasing use of advanced C^4I systems, information processing takes on new importance. Information processing is a component of decision-making in every aspect of planning, preparation, execution, and reconstitution. As military commanders make decisions, they have to repeatedly process updates on complex and dynamic battlefield situations, including staff estimates and higher headquarters directives. Digital information systems will almost certainly increase the amount and complexity of information provided to the staff, and training should provide the higher-order skills, such as information processing, needed on the digital battlefield.

Researchers have been examining the impact of increased information on decision-making, situational awareness, and team behaviors. Various models are being developed to explain how experts and teams might handle information-rich situations. Related developments include new ways to train and educate, and new ways to determine skill acquisition and proficient performance. This last issue, performance evaluation, is critical to training, but is often overlooked. More efficient and effective evaluation methods are needed to improve both training analysis and feedback mechanisms for training participants (Throne et al., 1999).

Three traditional methods of evaluation used in previous performance assessment research are: observation, survey, and interview (e.g., Throne et al., 1999). These methods have traditionally provided the measurement basis for conventional or pre-digital staff performance assessment (Crumley, 1989). Together, these methods can yield a multifaceted look at staff processes during training. A fourth method, examined here, is the use of automated measures of staff performance.

Digital technologies are uniquely suited to automatically collect user performance data. In fact, most computers routinely track all user inputs and system responses. Examples of user interactions that computers routinely log include "Back" keys, lists of recently opened files, and "Undo" commands. To achieve the training feedback and assessment potential of digital technologies, computer workstations must be more fully instrumented (Lickteig & Throne, 1999). Instrumentation refers to a log of all soldier-computer interactions correlated with the battlefield situation in which they occur. However, individual logs maintained on each operator's C^4I device are inadequate for examining collective performance by a group or staff of individuals working as a team. Fortunately, the collaborative nature of digital media readily

supports a network of C^4I devices and an integrated log of soldier-computer interactions across all members of a networked team. In fact, digital collaboration extends beyond the team to encompass other sensors and weapon systems that can be integrated into the network through common data links. Finally, a network of commanders, staffs, soldiers, weapon systems, and sensors can readily be linked to constructive or virtual training simulations to provide the stimulation to support staff training and performance assessment.

In response to the concerns and issues resulting from digitization, the U.S. Army Research Institute for the Behavioral and Social Sciences (ARI), Armored Forces Research Unit, is engaged in the design and development of training and performance evaluation techniques to support Force XXI digital capabilities (ARI, 1997). Research by ARI in this area seeks to capitalize on recent advances in the cognitive and behavioral sciences, and is focused on providing an empirical foundation for improved staff training and evaluation strategies for the digital battlefield of the future. This report details the assessment methodology work performed for a contract project titled, "Refinement of Methods for the Training and Assessment of Digital Staffs" and referred to herein as DC^4I-2. In support of that objective, a project team was developed, consisting of personnel from ARI, the Human Resources Research Organization (HumRRO) and Litton PRC (hereafter referred to collectively as the Project Team).

Specifically, the objective of the DC^4I-2 research project was to refine the staff and team training and assessment strategies that were designed during the original DC^4I project, "Prototype Methods for the Design and Evaluation of Training and Assessment of Digital Staffs and Crewmen." One of the lessons learned from the previous DC^4I research project was that additional research would be needed to determine whether the unit outcome of a staff action, processes of the staff, or a combination of both should be the focus of future automated measures development (Throne et al., 1999). This report concentrates on the refinement of command and staff assessment strategies, particularly automated measures of staff performance.[1]

The DC^4I-2 project began with a review of research literature and technical documentation related to team performance and assessment, with an emphasis on digital environments and automated performance data collection. A general design for staff performance assessment was formulated, based on findings of the literature review. An opportunity to implement this design was provided by an Army Concept Experimentation Program, the Battle Command Reengineering (BCR) IV, which took place in April, 2000. By participating in the BCR IV, the Project Team had the opportunity to conduct a trial implementation of the automated measures of performance. Coordination between ARI and the Mounted Maneuver Battlespace Lab (MMBL) at Fort Knox, Kentucky, enabled the two organizations to work together as a team to accomplish multiple goals. This report describes the development work of prototype automated measures, the results of their use during the BCR IV, and lessons learned for future staff performance assessment efforts.

[1] For more information on the refinement of the prototype team training package, see Deatz, Greene, Holden, Throne, and Lickteig (in preparation). For more information on the original DC^4I project, see Throne et al. (1999).

Organization of the Report

This report has five major sections:

- *Introduction:* Summary of previous research and relevant literature on team performance and evaluation. Includes a discussion of the experimental setting in which the project operated as well as a synthesis of previous findings on staff processes, measures of staff processes, and automated measures of staff performance.
- *Method:* Description of the procedure of front-end analysis, design, and development of automated measures of staff performance. Also includes a description of the BCR IV participants, BCR missions, and implementation of the developed measures.
- *Battle Command Reengineering IV Results and Discussion:* Representative results from developed automated measures implemented during BCR IV.
- *Battle Command Reengineering IV Implementation Lessons Learned:* Summary of the major lessons learned concerning evaluation and automated measures with implications for future MMBL experimentation.
- *Future Efforts*: Principles that should guide future research and development efforts on the use of automated measures to assess staff performance.

Appendix A contains a list of the acronyms and abbreviations used in this report. Appendix B contains the key aspects of the prototype automated measurement package, including operational definitions, rationale, and recommended output formats for the candidate automated measures and results additional to those presented in the Battle Command Reengineering IV Results and Discussion section. Finally, in addition to this report, the major research products associated with the training and evaluation for the BCR IV implementation are presented in the six-volume set of materials entitled *Training and Measurement Support Package, Battle Command Reengineering IV, Mounted Maneuver Battlespace Lab* (ARI, 2000). This set of materials is a revision of the package developed for the BCR III. The revised volumes pertaining to staff performance evaluation are:

- *Volume 5. Measures.* Includes copies of surveys and specifications for the automated measures used in BCR IV.
- *Volume 6. Data Codebook.* Contains basic descriptions and analyses of the variables in the data sets resulting from the data collection in BCR IV.

Battle Command Reengineering Setting

The unique training and measurement environment in which the Project Team operated had a significant impact on the direction and scope of the effort undertaken. The intent of this effort was to demonstrate prototype automated measures of staff performance in conjunction with the BCR Experimentation underway at Fort Knox. Surveys, observations, and interviews had already been developed for this environment and were not the focus of this project. This linkage to the BCR environment provided significant benefits, challenges, and some limitations that can be categorized into four broad areas: experimental objectives; surrogate command, control,

communications, and computer (SC^4) system; instrumentation and data collection and analysis; and participation of an existing, conventionally trained Army battalion staff. A discussion of each of these areas and their impact on the design of demonstrable prototype automated measures of staff performance follows.

Experimental Objectives

The objectives of BCR experimentation were to gain insights into how improvements in situational awareness and battlefield visualization affect battle command and to define requirements for future battle command C^4I systems. These objectives meshed very closely with the objectives of this project which are described earlier in this report. The BCR experiments seek to create future battlefield conditions that might exist in 2012 and beyond. These conditions include:

- The Army's information systems have been completely integrated, both vertically and horizontally.
- Expert computer systems will be performing routine information collection and dissemination tasks.
- Improved target and intelligence reporting capabilities will be seamlessly integrated into those information systems.
- Weapons systems will be lighter, more mobile, have increased range and lethality, and will be networked into the command and control (C^2) system.
- Robots will assume some functions that lend themselves to automation such as beyond line of sight battlefield reconnaissance and fuel and ammunition resupply.

The focal point for the BCR IV experiment was a futuristic conceptual battalion-level battle staff consisting of a commander and 13 staff officers and non-commissioned officers (NCOs). As shown in Figure 1, the battalion staff, which operated in a virtual simulation during the experiment, was configured into four nodes, associated with two battle command vehicles (BCVs) and two staff operations vehicles (SOVs). The node functions and job responsibilities for each staff member were left to the discretion of the battalion commander, who was allowed to reorganize the staff as he gained operational experience in the BCR environment. To encourage the staff to think "outside the box," few lessons learned and tactics, techniques, and procedures (TTPs) from previous BCR experiments were provided.

Subordinate company commanders and platoon leaders were located at constructive and virtual simulators. Company commanders used the Modular Semi-Automated Forces (ModSAF) constructive simulation, which took their inputs and used simulated personnel operating simulated combat systems to carry out their instructions. Platoon leaders and other combat vehicle crewmen were physically located in virtual simulators, which replicated the warfighting equipment of a postulated future ground fighting vehicle (FV).

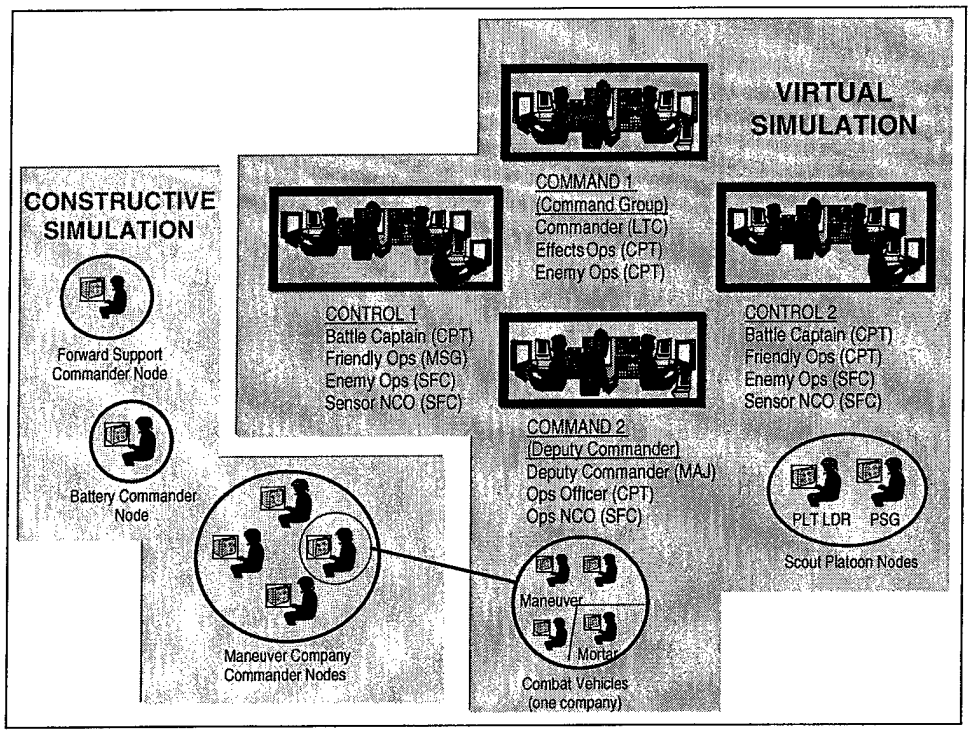


Figure 1. Battle Command Reengineering IV staff structure.

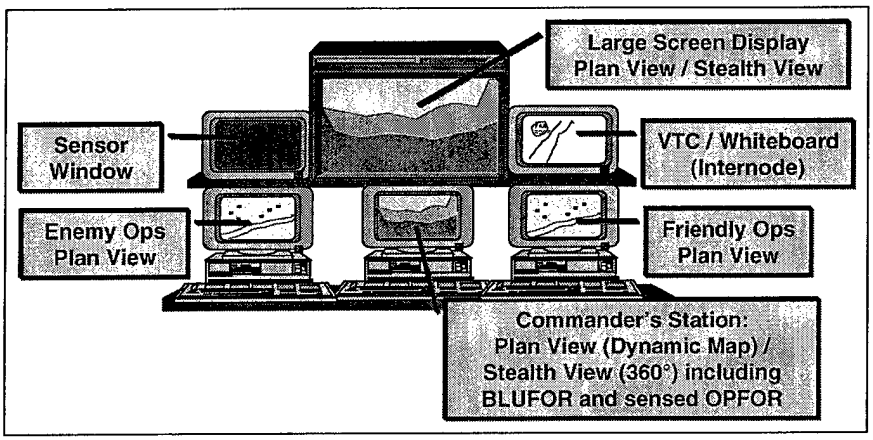
These conditions created several challenges for the Project Team in designing automated measures of staff performance. The first was the organization of the staff itself. There was no comparable battle staff organization with published standards of performance against which a measurement baseline could be established. Even within the BCR experimentation series, BCRs I-IV, changes in equipment and organization were made which precluded comparing the results from one experiment with another. Second, the role and functions associated with each staff member or node were not fixed. Even the physical location of a staff member could be changed between experiment trials or during a trial itself. Third, staff processes and staff products did not have to follow established Army doctrine or standing operating procedures (SOPs). The participants were encouraged to develop their own SOPs. As a result, the measures design had to be flexible enough to account for variations in staff processes and products.

Surrogate Command, Control, Communications, and Computers System

The SC^4 system provided the Project Team an opportunity to design and implement automated measures of staff performance for a functional C^2 system that had most of the characteristics envisioned for future C^4I systems. One of the challenges for the team was that, as an experimental system, the capabilities of the SC^4 and supporting systems are routinely modified to reflect the lessons learned from previous BCR experiments, and to include new features postulated to be available as information technology was improved. Also, the SC^4 system was not comparable to any currently fielded Army C^2 system so there was a training burden associated with using the system.

The SC^4 system used during BCR experiments is depicted in Figure 2 as it was configured in the nodes. The SC^4 system includes the following capabilities:

- C^2 Plan View Display (PVD), represented by the ModSAF two-dimensional PVD. On the PVD, the commander and the staff are able to view movements of all of their own systems, as well as any opposing forces (OPFOR) units detected by satellite or other sensors. Overlays can be drawn on the PVD, users can add labels or other notes, and there are tools that show past events and project future movements.
- Stealth display, providing a 3-dimensional visual representation of the battlefield with all of the systems that were visible on the PVD (i.e., friendly and detected OPFOR).
- Video teleconference (VTC) capability linking the commander and the staff.
- Collaborative Whiteboard capability, to allow the commander to present his intent and guidance to the staff visually and quickly. Users who are part of the Whiteboard session can show snapshots from their PVDs, draw in different colors on those images, add clipart-style labels and icons, and type words onto the Whiteboard.
- Large screen display, projecting a larger image of the PVD, Stealth, Whiteboard, or unmanned aerial vehicle (UAV) screens. The screen display is selectable by the node officer-in-charge and is meant to facilitate discussion about a particular display screen among the node members.
- Digitized Modified Combined Obstacle Overlay (MCOO), produced automatically for the large screen display, rather than as a manually produced intelligence overlay.
- Satellite imagery, acting as the electro-optic satellite sensor to deliver a direct downlink imagery feed.



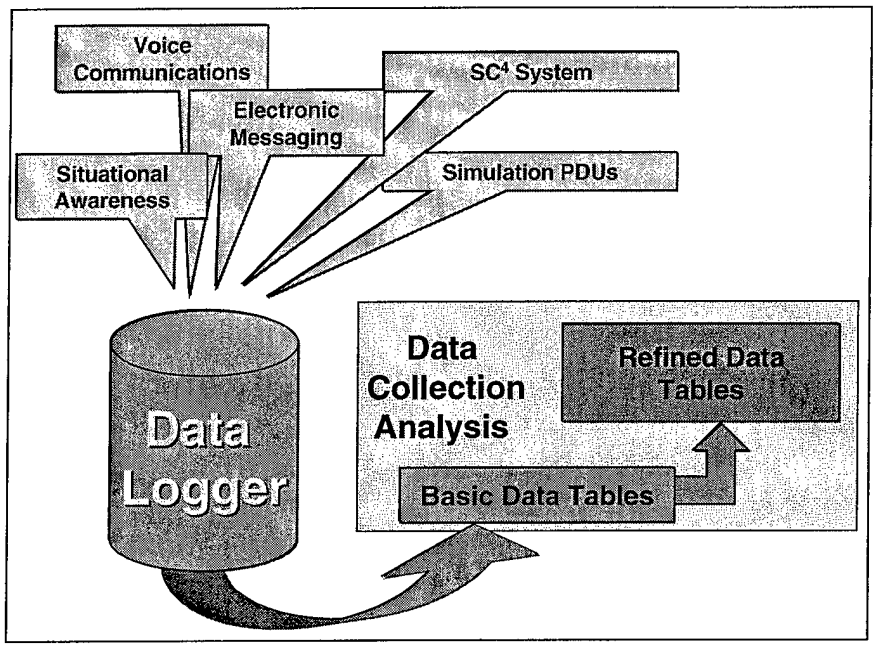
Note. VTC = video teleconference; Ops = operations; BLUFOR = blue forces; OPFOR = opposing forces.

Figure 2. Surrogate command, control, communications, and computers system setup.

Instrumentation and Data Collection

The ability to collect and integrate information about the usage of the SC^4 and communication systems and to relate that to simulated combat activities during BCR experiments was the catalyst to create automated measures of staff performance. The mechanism that creates this ability is the MMBL's Data Collection and Analysis System (DCA). Figure 3 illustrates key aspects and functions of the DCA. Information about SC^4 system usage,

electronic messaging, voice communications, displayed situational awareness, and the status of major combat systems in the constructive simulation driving the experiment are recorded by three separate systems, collectively referred to as the data logger.



Note. SC^4 = surrogate command, control, communications, and computers; PDU = protocol data units.

Figure 3. Battle Command Reengineering IV data collection system.

For BCR IV, major combat systems (vehicles and other weapon systems) were represented in either constructive or virtual simulation. Constructive simulation (ModSAF) was used to generate and control the OPFOR, friendly forces below the company level, and unmanned vehicles replicating both aerial and ground sensors (referred to as unmanned aerial vehicles [UAVs] and unmanned ground vehicles [UGVs], respectively). Constructive simulation workstations were used by the mortar battery commander, forward support company commander, four maneuver company commanders, and six deputy company commanders. The remainder of the experimental unit was in virtual simulation.

In the virtual environment, simulators were used to represent several vehicles. These included the battalion commander and deputy commander vehicles which were represented by the Advanced Research Projects Agency (ARPA) Reconfigurable Simulator Initiative (ARSI) simulator and an ARSI mockup, respectively; and BCVs and SOVs which were represented by command and control vehicle (C^2V) mockups. Scout vehicles and the manned platoon vehicles of one maneuver company were represented by Future Combat Vehicle mockups.

In the BCR setting, the constructive and virtual environments are linked by means of distributed interactive simulation (DIS) to form a seamless battlefield environment for the participants. Using protocol data units (PDUs), each major combat system sends information about its current status over the simulation network. This status is read by other combat systems

and by the SC^4 system. Before the SC^4 system displays and reports the information, it is filtered by a software program to present only that information that has been obtained by battlefield sensors, such as scouts, airborne radar, or combat vehicle crewmen, or by unit status reports. Additionally, this information is collected and processed by the DCA, which matches changes in combat system status to SC^4 system and communication system usage by time stamping each status change and SC^4 system use.

Using database extraction tools, the DCA initially creates a series of basic data tables. These tables can be subsequently refined into more advanced tables that answer specific questions posed by researchers. These refined tables were the starting point for the automated measures developed during this project. There are some challenges and limitations, however, to the DCA. The type of data available is limited by the content of the PDUs. One of the limitations of the system is that if desired information is not contained in a PDU, then it is not available for collection and analysis. Also, if the desired information is not identified prior to the start of an experiment, that information cannot be retrieved after the experiment has ended. Another limitation is the analysis of voice communications by staff members using tactical radios and vehicle intercoms. While these communications are recorded for later playback by researchers, there are no tools currently available in the Mounted Warfare Test Bed (MWTB) to automatically analyze the contents of the voice messages or to identify the participants. Finally, a major limitation of the current DCA is that it has a limited capability to produce graphs or charts. Any effort to expand this capability would require a major programming effort.

Army Battalion Staff Participation

The participation of an existing, conventionally trained Army battalion staff during the BCR IV experiment was viewed by the Project Team as a positive factor in designing and implementing automated measures of staff performance. While the staff would be operating in a new C^2 environment with unfamiliar equipment, weapons systems, and subordinate tactical units, they would also have established professional working relationships among themselves prior to the BCR IV experiment. The staff could focus on learning how to operate efficiently and effectively in the new environment without having the burden of learning how to work with one another. The challenge associated with this situation is that command and staff organizations have their own internal training and performance objectives which invariably differ from those required in the BCR environment.

Previous Findings

Based on previous research conducted to develop prototype training for future staffs (Throne et al., 1999), the Project Team determined that the current Army brigade and battalion level processes and products are radically transformed by the employment of systems like the SC^4. In fact, other researchers (e.g., Salas, Bowers, & Cannon-Bowers, 1995) have similarly concluded that the introduction of automated systems will change staff members' roles and responsibilities and that research is needed to identify the changes in staff processes that are necessary to perform effectively with these systems.

As described earlier, with no published Army doctrine for future digital staff processes or staff products applicable to the BCR IV setting that could form the basis for automated measures

design for this project, the Project Team turned to the research literature for automated measures of military staff performance, especially at the brigade and battalion level. Unfortunately, there was little available information applicable to the project's objective. There was, however, a considerable amount of information about teams and team processes that could guide a measures design and development process. Based on the definition of teams provided below, a military staff could be viewed as a team. Therefore, the Project Team followed this approach during the project and considered staffs as a type of team.

In order to develop effective automated C^2 team performance measures, the term "team" had to be defined. A relatively comprehensive definition of team was developed by Salas, Dickinson, Converse, and Tannenbaum (1992). They define team as "a distinguishable set of two or more people who interact dynamically, interdependently, and adaptively toward a common and valued goal/objective/mission, who have each been assigned specific roles or functions to perform, and who have a limited life-span of membership" (p. 4). As indicated by Baker and Salas (1997), this definition of team has led to a common understanding among researchers of the aspects that define a team. For purposes of measurement, this common understanding is important because it establishes the boundaries of what constitutes a team and it also "defines...team processes, and team outcomes that should be accounted for in the measurement process. Essentially, it tells us what to measure when assessing team performance and presents a foundation on which to construct team performance measurement tools" (Baker & Salas, pp. 332-333).

Researchers interested in developing team performance measures need to distinguish between team processes, team products and unit outcomes as well as their interrelationships (Cannon-Bowers & Salas, 1997). Processes are the activities, responses, and behaviors that people use to accomplish tasks. Examples in a military staff context include decision-making, orders production, monitoring, and synchronization (Department of the Army [DA], 1997a). Team products are the results of the team processes. Examples of military staff products include situational awareness information, conclusions, recommendations, guidance, intent, concept statements, overlays and orders (DA, 1997a). Staff processes and products can be equated to measures of performance, while unit outcomes can be equated to measures of effectiveness. Outcomes or products, unlike processes, are usually not diagnostic because they do not lead to an understanding of the underlying causes of performance or how performance could be improved (Smith-Jentsch, Johnston, & Payne, 1998).

In both military and business settings, outcome measures are more popular since they are easier to obtain and seem to be more objective (Smith-Jentsch et al., 1998). However, outcome measures, by themselves may hinder learning and cannot facilitate remedial training (Johnston, Smith-Jentsch, & Cannon-Bowers, 1997). For example, if a team reaches the correct decision through a faulty process or by luck, outcome measures would reinforce a process that might result in incorrect decisions at a later time. Since outcome measures alone cannot determine training needs and provide useful feedback to the trainees, team processes need to be linked to outcomes for a comprehensive understanding of team performance (Smith-Jentsch et al., 1998).

A model for developing team and individual performance measures that includes both process and outcome measures is provided by Cannon-Bowers and Salas (1997). According to

their model, team process measures can be obtained through observational scales, expert ratings, content analysis, and protocol analysis, while team outcome measures can be obtained through observational scales, expert ratings, critical incidents, and automated performance recording. This last method, automated performance recording, has been used primarily for measuring individual performance, but is rarely used for evaluating team processes. It has the potential for providing objective data and instant feedback with little or no intrusion on training participants. This possibility for improving staff performance measurement by means of automated data collection is at the core of this effort.

For this project, three major aspects of performance measurement were examined within the military, academic, and business literature: team and staff processes, measures of team and staff processes, and automated measures of performance. First, various staff processes that might affect performance outcomes were examined to find processes that could be at least partially measured through automated means. Many brigade- and battalion-level staff processes are verbal. The DCA cannot analyze the content of voice communications, hence automatically measuring voice processes is not currently feasible. However, the Project Team hypothesized that staff processes might be automatically assessed if suitable measures of performance could be developed. Second, since the focus was on staff training in a digital environment, the literature on various measures of staff performance in information-rich C^2 environments was examined. Finally, literature on automated methods for obtaining team process data, both to assess training quality and to provide feedback, was explored.

Staff Processes

As indicated earlier, defining team processes could lead to determining what to measure (Baker & Salas, 1997). Therefore, the Project Team looked at several models of team processes that could form the basis for developing measures of staff performance. One of the earliest classifications of staff processes was developed by Olmstead (1992). His seven processes (referred to as organizational processes) include: sensing, communicating information, decision-making, stabilizing, communicating implementation, coping actions, and feedback. However, Olmstead's processes are more related to task work rather than teamwork (Sterling & Quinkert, 1998). Sterling and Quinkert define task work as individual activities that are usually technical in nature, such as analyzing information, making decisions, and communicating orders, while they define teamwork as interpersonal activities that enhance communications and interaction patterns, such as coordination of activities.

Another model of teamwork, developed by Brannick, Prince, Prince, and Salas (1995), assessed teamwork based on skill dimensions developed through a review of the literature and prior research. Their dimensions included: adaptability/flexibility, situational awareness, leadership, assertiveness, communication, and decision-making/mission analysis. In an earlier study, Baker and Salas (1992) included both giving suggestions or criticisms and coordination in their teamwork skill dimensions as well as some of the dimensions mentioned in Brannick et al. Additionally, in a review of the U.S. Navy research literature on tactical teams, McIntyre and Salas (1995) describe the Critical Team Behaviors Form, which was designed to identify effective and ineffective team behaviors. Critical behaviors were categorized into seven dimensions: adaptability, coordination, communication, acceptance of suggestions or criticism, giving of suggestions or criticism, team spirit and morale, and cooperation.

One of the most thorough evaluations of team processes was conducted by Cannon-Bowers, Tannenbaum, Salas, and Volpe (1995). These authors conducted an extensive review of the literature related to defining teamwork skills, grouped similar skills together under a single skill dimension, and developed a common definition for each dimension that emerged. They concluded that there is a core group of eight skill dimensions common to most team-based tasks. The definitions given by Cannon-Bowers et al. are provided in Table 1.

Table 1

Cannon-Bowers et al.'s Team Process Skill Dimension Definitions

Team Process Skill Dimension	Definition
Adaptability	The process by which a team is able to use information gathered from the task environment to adjust strategies through the use of compensatory behavior and reallocation of intrateam resources.
Performance Monitoring and Feedback	The ability of team members to give, seek, and receive task-clarifying feedback; includes the ability to accurately monitor the performance of teammates, provide constructive feedback regarding errors, and offer advice for improving performance.
Leadership/Team Management	The ability to direct and coordinate the activities of other team members, assess team performance, assign tasks, motivate team members, plan and organize, and establish a positive atmosphere.
Interpersonal Relations	The ability to optimize the quality of team members' interactions through resolution of dissent, utilization of cooperative behaviors, or use of motivational reinforcing statements.
Shared Situational Awareness	The process by which team members develop compatible models of the team's internal and external environment; includes skill in arriving at a common understanding of the situation and applying appropriate task strategies.
Communication	The process by which information is clearly and accurately exchanged between two or more team members in the prescribed manner and with proper terminology; the ability to clarify or acknowledge the receipt of information.

(table continues)

Table 1 (Continued)

Team Process Skill Dimension	Definition
Coordination	The process by which team resources, activities, and responses are organized to ensure that tasks are integrated, synchronized, and completed within established temporal constraints.
Decision-Making	The ability to gather and integrate information, use sound judgment, identify alternatives, select the best solution, and evaluate the consequences.

Note. Adapted from Cannon-Bowers et al. (1995), pp. 344-346.

A taxonomy of team processes similar to that of Cannon-Bowers et al. (1995) is presented in U.S. Army Training and Doctrine Command (TRADOC) Regulation 350-70 (DA, 1999). According to this regulation, teams must possess both team knowledge and team skills to perform as part of an effective team. Team knowledge requirements include the nature of teamwork and teamwork skills, the importance of teamwork, and team member responsibilities. Team skill requirements include communication, cooperation, building on other team members' strengths and weaknesses, adaptability, shared situational awareness, coordination, providing opinions and ideas for improvement, conflict resolution, decision-making, and team leadership. These team skill requirements are very similar to Cannon-Bowers et al.'s team process skill dimensions. Table 2 provides a summary of the models reviewed and their relation to Cannon-Bowers et al.'s classification scheme of team process skill dimensions.

Table 2

Classification of Skill Dimensions by Teamwork Model

Skill Dimensions[a]	Authors				
	Baker & Salas, 1992	Olmstead, 1992	Brannick et al., 1995	McIntyre & Salas, 1995	TRADOC (DA, 1999)
Adaptability	✓		✓	✓	✓
Feedback[b]	✓	✓		✓	✓
Shared SA			✓		✓
Leadership			✓		✓
Interpersonal Relations[c]	✓			✓	✓
Communication	✓	✓	✓	✓	✓
Coordination	✓			✓	✓
Decision-Making		✓	✓		✓

Note. DA = Department of the Army; SA = situational awareness; TRADOC = U.S. Army Training and Doctrine Command.

[a]from Cannon-Bowers et al. (1995)

[b]Feedback is also referred to as acceptance of/giving suggestions, criticism.

[c]Interpersonal relations is also referred to as cooperation, conflict resolution, assertiveness, and morale building.

As indicated in Table 2, it appears that most of the models reviewed contain similar or overlapping dimensions. The dimensions outlined by Cannon-Bowers et al. (1995) are quite likely the most comprehensive and inclusive since these researchers considered most of the other research mentioned in this section when developing their team process skill dimensions. The only model that has any dimensions not included in the Cannon-Bowers et al. model is Olmstead's (1992), which as mentioned previously, measures many task work skills in addition to some teamwork skills. The TRADOC (DA, 1999) model's skill requirements are essentially the same as Cannon-Bowers et al.'s team process skill dimensions, although the TRADOC model does not provide the same degree of background information about team process skills as does the Cannon-Bowers et al set. The Cannon-Bowers et al. model provides detailed information about the research that led to the formulation of their team skill dimensions. Therefore, the Project Team adopted the Cannon-Bowers et al. (1995) model of team process skill dimensions as a starting point for the automated measures development.

Measures of Staff Processes

As mentioned previously, this part of the review will focus on research that has integrated and objectively evaluated both staff processes and products rather than research that has only evaluated unit outcomes. In a thorough review of Army C^2 performance, Crumley (1989) states that one of the most frequent problems researchers face is what criteria determine effective C^2 performance. There are two basic approaches that can be used to solve this problem: measures based on battle outcome (" 'It's not how you play the game, it's if you win or lose that counts' " [p. 100]) or measures based on the command post as an entity (" 'It's not if you win or lose, it's how you play the game that counts' " [p. 100]).

After identifying and relating staff process and outcome measures, researchers must identify the assessment methods. The two general methods of assessment are subjective and objective assessment, both of which may be appropriate for team process skills. Most feedback to the Army centers and schools is based on subjective or qualitative data, which includes informal comments, surveys, and interviews (Mohs, MacDiarmid, & Andrews, 1988). Objective or quantitative data, which Mohs et al. define as data derived from observation, analysis, and testing of performance, are used less frequently due to availability and cost.

However, there are several difficulties associated with assessing team performance. Cooper, Shiflett, Korotkin, and Fleishman (1984) cite four: a) evaluation of teams rather than single individuals; b) behaviors that are difficult to observe and evaluate; c) complexity of environment; and d) no standard criteria by which to evaluate performance. A fifth difficulty relates to observer training and staffing. It can be very time-consuming and expensive to train observers on how to assess performance on certain tasks. However, as Dwyer, Fowlkes, Oser, Salas, and Lane (1997) mention, subject matter experts (SMEs) need to be highly trained in order to achieve even moderate levels of interrater reliability (i.e., consistent judgments among raters), and consequently, objective assessment. This is because observers have trouble discriminating among multiple skill dimensions and they can also be prone to decision biases when making judgments (Smith-Jentsch et al., 1998).

Other problems with observation as an objective form of measurement have been identified by Dwyer et al. (1997). For example, they propose that one of the major problems with the

assessment of C^2 team performance is that typically the evaluation is based on mostly numerical ratings of global performance provided by SMEs. Even though these types of ratings can be useful in assessing overall performance, the authors point out that they have some limitations. For example, since the ratings are usually global in nature, they lack diagnostic specificity. Good performance during one part of training may be cancelled out by poor performance during another part. Secondly, even though observation usually requires highly trained SMEs, they still only reach low levels of interrater reliability. Therefore, even observations can be subjective, particularly if the tasks to be observed are not clearly delineated and defined.

The major problem associated with observations when they are subjective is that although subjective data can be useful, the utility will be suspect if the data cannot be shown to be reliable and valid. Reliability refers to the consistency and stability of the measure over time or measurement points, while validity refers to the degree to which the measure achieves its goal and measures what it is meant to assess. As Mohs et al. (1988) state, "reliable and valid measures of performance are necessary to determine if trainees have acquired the skills intended..." (p. 14). If observers are not highly trained, then they most likely will attain low levels of interrater reliability, which means their results will not be reliable and consequently, not valid. Therefore, even though subjective measures typically require less cost and effort, objective measures have a greater chance of being not only reliable but also valid.

Additionally, Mohs et al. (1988) claim that there is too much reliance on subjective performance measures, especially in the Army. In their review of the literature, Baker and Salas (1992) concluded that researchers rely heavily on observation to measure team processes. The authors suggest that "critical skill dimensions and behaviors" (p. 472) need to be identified in order to achieve high interrater reliabilities. In the past few years, researchers have attempted to follow Baker and Salas' suggestion and develop more standardized and objective observation methods to improve evaluation procedures. Three of the more promising approaches are described below.

For example, Hayes, Layton, and Ross (1995) developed observation-based outcome measures for the Army Command and Control Evaluation System (ACCES) for seven categories: general, incoming information handling, situation assessment, course of action analysis, preparation of directive, outgoing information handling, and information exchange. The authors point out that ACCES does not provide a rating of C^2, only indicators of its effectiveness. Moreover, ACCES cannot provide rapid feedback in the training environment since the quantitative data (e.g., graded quizzes, observer journals and data collection forms, participant interviews) take several days to process. However, once the data are analyzed, insights can be provided.

Another attempt to produce standardized observation methods is provided by Fowlkes, Lane, Salas, Franz, and Oser (1994). They define critical behaviors to measure aircrew coordination based on the targeted acceptable responses to generated events or tasks (TARGETs) methodology. The goal was to develop a measurement tool that could capture team processes, identify performance deficiencies, and produce a reliable method for monitoring training progression. The measurement tool also needed to be sensitive to the training constructs of interest, be responsive to measurement in applied team training contexts, and minimize rater bias

and error. The key aspects of the TARGETs methodology are that it is event-based with acceptable task responses identified for each event beforehand, the behaviors of interest are controlled by the selection of events presented as well as when they are presented, and there are specific behaviors that observers without detailed subject matter expertise can identify. According to the researchers, the TARGETs approach accomplishes all the goals mentioned and it "provides a capability to evaluate a variety of team member behaviors in a manner that is psychometrically sound" (p. 60).

Finally, Dzindolet et al. (1998) designed a study to determine the extent to which certain exercises helped train the Officer Basic Course students. This study identified ways to measure performance on the learning objectives, which were identified by Dzindolet et al. in previous work. Three different types of data were collected: objective process measures (e.g., accuracy of battle tracking, efficiency of spot reporting), subjective process measures (self-reported proficiency), and objective outcome measures (multiple choice tests). However, the researchers encountered many problems with each of the objective process measures. First, some of the measures were distorted because of the simulation characteristics. Second, some instructors took a more active role than others in helping students when they made mistakes and this variability made it difficult to examine performance. Third, due to level of experience and training, the ability levels of the instructors varied.

Based on their problems with developing the measures, Dzindolet et al. (1998) maintain that until objective process measures are improved, researchers will have to use self-reports and subjective ratings. Suggestions provided by the authors on how to improve objective process measures include allowing simulation developers to work with instructors and researchers to eliminate the distorting simulation characteristics, providing instructors with a training program that would identify appropriate instructor behaviors and levels of intervention, and giving instructors practice time so they have comparable ability levels.

Traditional methods for assessing C^2 performance, including staff performance, are heavily "burdened" measures (i.e., resource intensive). The actual measurement of staff performance is a task almost invariably performed by human observers. The measurement burden begins with the multiple roles and responsibilities assigned to observer personnel, generally referred to as observers/controllers (O/Cs). The multiple tasks required by O/Cs, in addition to observation and data collection, are only suggested by their dual title. In fact, an analysis of O/C task responsibilities provided by Brown, Nordyke, Gerlock, Begley, and Meliza (1998) underscores the measurement burden: "However, the analysis clearly shows that O/Cs are involved in data collection tasks which divert them from player behavioral observations, coaching and mentoring" (p. 52).

The measurement burden is further compounded by the inherent complexity of C^2 performance including the requirement to track and integrate communication exchanges across numerous and dispersed participants (Crumley, 1989; Olmstead 1992). As a result, traditional staff performance measurement methods (e.g., surveys and observations) often fail to provide a detailed and objective database for adequately assessing staff performance. Moreover, the workload and time lags required for manual integration and interpretation of staff performance

data obtained by traditional methods often prevent timely and meaningful feedback to the training audience.

As can be seen from the research reviewed, most C^2 process measures are obtained through observational techniques that are often neither reliable nor valid. Perhaps that is why people tend to trust objective, quantitative data, especially if it is automated or computer-generated (Smith-Jentsch et al., 1998). Finally, even though most of the research reviewed has tried to standardize its subjective measures, an observer is still required to obtain these measures, which translates into more time and effort to train the observers. Therefore, more efficient, effective, and objective process measures of C^2 team performance are needed. One of this project's main objectives was to meet this goal by expanding the research on methods of obtaining objective team performance assessment through automated data collection.

Automated Measures of Staff Performance

Automated measures, as used in this report, are initially generated by DCA computer software programs or routines resident in the MWTB. These programs automatically collect, reduce, compile, format, and display soldier and machine interaction performance data. Some of the advantages staff trainers and evaluators can expect to gain by employing automated measures of staff performance are provided below.

First, the more staffs rely on computers to do their work, the more their computer interactions will become meaningful aspects of work process and outcome. Second, automated measures are not only objective, they are direct measures *of* performance. In contrast, many traditional measurement methods (e.g., survey, interview, observation) are measures *about* performance. Third, a greater reliance on automated measures may increase the scope and precision of performance assessment and feedback. Fourth, automated measurement and analysis may be needed for more complex work settings, such as C^2 staff performance. Finally, unobtrusive and automatic data collection may reduce measurement error as well as observer workload and resource requirements.

In reviewing the available literature, the Project Team found very few studies of objective automated measures of staff performance. Therefore, this section of the literature review is quite brief since it focuses only on studies that have developed automated process measures of team performance in C^2 environments.

Atwood et al. (1991) used automated measures to collect data on training of company-level simulations on the Combat Vehicle Command and Control (CVCC) systems. Data were collected and analyzed using the DCA. The DCA is a set of tools designed to collect, reduce and analyze battlefield performance, C^2, communications, and other types of data in distributed simulations. The automated outcome measures were used to partially assess the training outcome of preparing the soldiers to use the equipment. However, since the study was focused on the evaluation of the training program rather than the measures used to evaluate it, no conclusions are presented on the utility of the automated measures.

Many other studies testing the CVCC system and its capabilities have used automated measures gathered through the DCA (e.g., Leibrecht, Meade, Schmidt, Doherty, & Lickteig, 1994; Lickteig & Collins, 1995). These studies gathered data on issues relating to the battlefield operating system (BOS) functions of C^2, maneuver, fire support, and intelligence. For example, a wide range of automated measures were developed by Leibrecht et al. (1994) to assess training, soldier - machine interactions, and operational effectiveness. Figure 4 provides a sample of these automated measures collected for the C^2 BOS.

Sample of Command and Control Automated Performance Measures from Combat Vehicle Command and Control Battalion Evaluation

Receive and Transmit Mission

Mean time for transmission of fragment order (FRAGO) across echelons (battalion, company, platoon)
Duration of requests by company and platoon members to clarify FRAGO/overlay*
Consistency of FRAGO received
Number of requests by company and platoon members to clarify FRAGO/overlay*

Receive and Transmit Enemy Information

Time to transmit Intelligence Report (INTEL) across echelons
Consistency of INTEL received across echelons
Number of requests to clarify INTELs by company and platoon members*

Receive and Transmit Friendly Troop Information

Mean time to transmit Situation Report (SITREP) across echelons
Mean duration of voice transmissions within and between echelons
Deviation of BLUFOR locations reported in SITREP from actual location
Delay between observed phase line or line of departure crossing and reported crossing
Delay between observed battle position arrival and reporting SET at battle position
Elapsed time from request for fuel and/or ammunition report across echelons
Number of voice transmissions within and between echelons

Manage Means Of Communicating Information

Average length of voice radio transmissions by echelon
Total number of voice radio transmissions by echelon
Total time on voice radio network by echelon
Number of named voice reports by echelon*

* Indicates manual measures extracted from voice data recordings.

Figure 4. Sample command and control battlefield operating system measures.

Finally, for the original DC^4I project, Throne et al. (1999) designed and tested a set of 14 prototype automated measures of staff performance. These automated measures were an initial attempt at measuring processes through automated means to enhance performance feedback to future battle staffs. The goal was to gain experience with acquiring information from various sources, including digital information from an SC^4 system. Automated measures were designed to answer some of the MMBL's BCR III issues focusing on staff performance. Table 3 provides

the subset of the BCR III issues for which the DC^4I team developed the automated measures. Observer and survey measures were also developed to complement the automated measures and ensure that the MMBL's issues could be addressed if the automated measures did not produce the anticipated results. A complete set of the BCR III evaluation issues and methods is provided in the MMBL's Experiment Plan (MMBL, 1998). Table 3 also contains descriptions of all the automated measures designed for the issues.

Table 3

Reengineered Battle Command (RBC) III Issues and Automated Measures of Performance

Issue	Measure of Performance	Description
1. Can the RBC decrease the time for planning and increase the time to prepare and rehearse?	Battalion staff communication patterns	Total time spent using each of the communication tools (e.g., e-mail, Whiteboard) for each node position for each planning and execution session.
2. Can the RBC provide the information and support system to assist the commander's decision-making process?	a. UAV mission effectiveness	Number of OPFOR vehicles first detected by UAVs, divided by total number of OPFOR vehicles detected.
	b. Percent of enemy vehicles inside the battalion's area of responsibility that were detected	Number of unique OPFOR vehicles detected by sensors, scouts, or weapons systems controlled by the battalion, divided by the number of OPFOR vehicles.
	c. Percent of enemy vehicles inside the battalion's area of responsibility that were never detected	Number of unique OPFOR vehicles not detected by sensors, scouts, or weapons systems controlled by the battalion, divided by the number of OPFOR vehicles.
	d. Use of SC^4 communication tools during mission	Total time spent using each communication tool and the number of communication tool initiations per mission for each node position.
	e. Use of SC^4 tools allowing analysis of past, present, and projected battlefield positions	Total time spent using each tool and the number of tool initiations for each node.
3. Can the RBC allow efficient synchronization of combat, combat support, and combat service support assets?	a. Percent of OPFOR kills inside effects box	Number of OPFOR kills in effects box, divided by total number of OPFOR kills.
	b. Ratio of indirect to direct fire OPFOR kills	Ratio of number of indirect fire kills to number of direct fire kills.

(table continues)

Table 3 (Continued)

Issue	Measure of Performance	Description
4. Does the RBC provide efficient battle tracking and facilitate precise execution?	a. Percent of OPFOR vehicles engaged from flank or rear	Total number of flank or rear engagements on OPFOR vehicles, divided by total number of OPFOR vehicles.
	b. Percent of friendly force vehicles engaged from flank or rear	Total number of flank or rear engagements on friendly force vehicles, divided by total number of friendly force vehicles.
	c. Average range of OPFOR fire engagements	Average range of friendly weapon systems, by type, against OPFOR vehicles that were killed by mission.
5. Does the RBC contribute to more rapid and efficient destruction of enemy forces?	a. OPFOR vehicle kills by friendly weapons types	Number of OPFOR vehicle kills by each friendly weapon type by mission.
	b. Time to destroy OPFOR	Time from first OPFOR engagement until OPFOR vehicle losses exceed 70%.
6. Can the RBC increase the span of control of the commander?	Number of subordinate unit leaders commander personally contacted during mission execution	Commander's frequency of and amount of time for use of communications tools across each of the different personnel with which he interacted during mission execution.

Note. OPFOR = opposing forces; SC^4 = surrogate command, control, communications, and computers; UAV = unmanned aerial vehicle.

For example, to assess the SC^4 system's capability to support the commander's decision-making process, five automated measures were implemented.[2] One of these focused on the performance of the UAVs, which were controlled by the battalion staff, and micro-UAVs (MUAVs), which were controlled by the battalion scout platoon (Table 3, Measure 2a). These two systems, properly employed, could provide the commander with accurate, timely information about the location and activity of the OPFOR. Over the course of the nine mission trials during the BCR III, the UAVs and the MUAVs were the first to detect 51% of all OPFOR vehicles in the battalion's area of interest.[3] In a further breakdown, the two systems detected 32% of tanks, 53% of infantry fighting vehicles (IFV), and 56% of artillery vehicles. When all types of sensors (i.e., ground, airborne, and satellite) were included (Table 3, Measure 2b), 70% of OPFOR vehicles were detected.

[2] A complete description of the results of the automated measures developed for BCR III can be found in the *Training and Measurement Support Package, Battle Command Reengineering III, Mounted Maneuver Battlespace Lab* (ARI, 1999).

[3] The battalion's area of interest extended 15 kilometers beyond the unit's area of responsibility. The unit's area of responsibility was delineated by its rear, flank, and forward boundaries assigned by its higher headquarters (DA, 1997b).

Notably, the majority of these vehicles were first detected in the area of interest, beyond the area of operations, which gave the commander more time to assess OPFOR strength, capabilities, and intentions before they closed to within range of the unit's combat systems. For the 30% of OPFOR vehicles not detected, the majority were second echelon tanks or artillery systems (Table 3, Measure 2c). None of these undetected systems had an impact on the outcome of any mission. Grouping the data results from these three measures together, it could be demonstrated that the commander knew the location of 70% of the OPFOR vehicles in sufficient time to make an informed decision.

This example illustrates how valid and reliable data, generated automatically by the DCA, were used to address experimental issues related to staff processes. Overall, results from all of the automated measures developed under this project helped the MMBL conclude that the reengineered battle command provided the information and support system needed to assist the commander's decision-making process. More complete results for the sample of automated measures described as well as the other automated measures listed in Table 3 are available in the MMBL's Experiment Final Report (1999), or in Throne et al. (1999).

Summary

The Army's growing reliance on C^4I systems underscores the need for additional research and development on automated measures based on the ability of digital technology to support training assessment and feedback. In particular, automated measures are needed that address staff performance and the expanding role of C^2 systems for information-age forces. A key challenge is how to obtain automated measures of staff processes and products. To date, much research has been conducted on teams and the processes that make a successful team. This project was an attempt to implement some of the findings of the previous research in the development of automated measures to assess staff processes.

Method

The BCR IV provided the context in which the automated measures of staff performance for this project were designed, developed, and implemented. This section of the report details the front-end analysis, design, and development of the automated measures as well as the unique characteristics of the BCR environment which influenced all aspects of automated measures implementation.

Measures Development Process

The goal of this effort was to expand the automated measures development work initiated in BCR III for implementation in future C^2 experimentation. The intent was to explore and implement some of the data processing capabilities of advanced C^4I systems and military simulations to develop a set of automated measures. These automated measures could potentially support staff performance, provide immediate individual and staff/team feedback, reduce the need for outside expert observers, facilitate after action review (AAR) preparation and delivery (with formatted tables and graphs from the DCA), and facilitate improvement of the staff information exchange process. A description of the method used to determine which staff

processes should be measured is presented below. The steps in the staff process measures design and development are also described.

Front-End Analysis

The front-end analysis began with an examination of the eight team process skill dimensions identified in the Cannon-Bowers et al. (1995) model. As discussed earlier, the Cannon-Bowers et al. team process skill dimensions were selected as the organizational framework from which to develop the automated measures. These processes are quite similar to those identified by TRADOC (DA, 1999) and Army leaders and trainers should already be familiar with most of the terminology. Therefore, this framework should be useful as a way of presenting team performance feedback in a military setting.

The Project Team initially focused on which of the eight skill dimensions might be, at least partially, addressed by automated measures. Two of the eight dimensions were not selected: leadership/team management and interpersonal relations. The Project Team considered leadership to be more of an individual skill rather than a team skill. Since this project was more interested in staff processes, leadership was not addressed in the measures of performance. Also, interpersonal relations were not measured because most adequate measures would require an observer and could not be measured through automated means. Therefore, six of the eight skill dimensions identified by Cannon-Bowers et al. were selected by the Project Team as potential candidates for being partially measured through automated means. The six dimensions chosen for automated measurement were: Adaptability, Performance Monitoring and Feedback, Shared Situation Awareness, Communication, Coordination, and Decision-Making.

After identifying the skill dimensions that could be partially measured, the Project Team then determined which supporting team-related behaviors could be measured using DCA data. After reviewing all of the team process models previously cited, a comprehensive list of behaviors was created. The behaviors associated with each skill dimension that were considered by the Project Team are identified in Table 4. The behaviors thought to be at least partially measurable by the DCA data are italicized in the table.

Table 4

Team Process Skill Dimensions and Associated Behaviors

Adaptability
Have back-up plans[a]
Transition smoothly to back-up[a]
Adjust quickly to situational change[a]
Perform tasks outside job when asked[c]
Change way task performed when asked[c]
Adapt to information presented out of order[c]
Performance Monitoring and Feedback
Respond to requests for information[d]
Accept reasonable suggestions[c,d]
Avoid nonconstructive comments[c,d]
Provide constructive suggestions[d]
Ask for advice when needed[c,d]
Ask for input regarding performance[c,d]
Observe and keep track of team members' performance[d]
Recognize when team members make mistakes or perform exceptionally well[c,d]
Listen to other team member communications[d]
Obtain information about the outcomes of decisions and actions where appropriate[b]
Modify activities or make new plans or decisions based on follow-up information[b]
Shared Situational Awareness
Maintain the big picture[a]
Identify potential or anticipated problems[a,d]
Remain aware of resources available[a]
Provide information in advance[a]
Note deviations from steady state[d]
Avoid tunnel vision and maintain awareness of all relevant contacts[d]
Recognize others' information needs[d]
Recognize need for action[d]
Be proactive[d]
Communication
Use correct format[a,c]
Use proper terminology[a,c,d]
Maintain clarity in communications[a,c]
Provide acknowledgments[a,d]
Do not engage in excess chatter[c,d]
Report all contacts or detections as prescribed[d]
Pass complete information to correct members[b,d]
Provide information in complete, accurate, timely, and efficient manner[b]

(table continues)

Table 4 (Continued)

Coordination
Synchronize actions[a,d]
Pass relevant information in timely and efficient manner[a,c,d]
Be familiar with others' jobs[a,c,d]
Facilitate performance of other team members[c,d]
Avoid distractions during critical operations[d]
Decision-Making
Use all relevant, available information[b]
Make mostly correct decisions based on information available[b]
Make decisions in timely manner[b]

[a]Dwyer, Oser, Salas, & Fowlkes, 1999
[b]Olmstead, 1992
[c]Oser, McCallum, Salas, & Morgan, 1989
[d]Smith-Jentsch et al., 1998
Note. Behaviors thought to be measurable by the Data Collection and Analysis System appear in italics.

Next, the lessons learned during BCR III (Throne et al., 1999) were reviewed for guidance on how the existing prototype automated measures could be improved for BCR IV. One of the key lessons was that more collaboration is needed among the researchers, SMEs, and programmers developing the measures. For example, operational definitions may need to be revised to specify the format in which the automated measures data is to be reported. This could allow the programmer extracting data to configure the output so it can be readily converted into a format suitable for staff performance feedback sessions. The DCA provides the programmer with a multitude of methods to extract, manipulate, and format data. Working with both the operational definition of an automated measure and the specific format in which it is to be reported, the programmer will be better able to meet the requirements for training and performance assessment. In order to increase collaboration and to ensure the measures developed were supporting the various BCR IV research objectives as much as possible, a SME Advisory Group (SAG) was formed to evaluate the measures recommended for development by the Project Team. Members of the SAG included representatives from the MMBL and their supporting contractors as well as the Project Team.

A second lesson learned was that measures under development normally require iterative data processing sessions or analytic "runs" to refine them. The initial runs may provide insights into reformatting the output or limiting the parameters to exclude extraneous information. Preliminary analysis of the data could also point to another area to be investigated that was not considered during the analysis and design of the measures. Therefore, the Project Team and MMBL programmers agreed to use the prototype automated measures for a single BCR mission to analyze the measures output. If the outputs were satisfactory, then all the data would be included in the subsequent analysis. However, if the outputs were not as expected, the operational definitions and output formats would be refined to improve the measures.

A third lesson learned was that there was a need to focus the data collection and analysis effort. For that reason, for BCR III, an exercise flag log was created which was used by

observers to record the timing of key events during a mission, such as the receipt of a higher headquarters order or crossing the line of departure. At the conclusion of BCR III, the Project Team found that the complexity of the BCR environment (e.g., dispersed staff; multiple methods of communication, such as e-mail, radio, and Whiteboard conferences; and simultaneous planning and execution cycles) led to significant errors in manually maintaining the exercise logs. Consequently, the BCR III Project Team recommended that a method be found to automate these exercise logs.

Design

The first step in the design process was to select candidate measures for development. For each of the six team process skill dimensions chosen, multiple candidate measures were identified in order to obtain supporting data on that particular skill dimension. Table 5 shows the candidate measures identified for each process skill dimension.

Table 5

Candidate Automated Measures

Team Process Skill Dimension	Candidate Measure of Performance
Adaptability	Terrain Analysis
	Node Location
	Loss of Node
Performance Monitoring and Feedback	Situation Report (SITREP) Use
	Spot Report (SPOTREP) Use
	Commander's Critical Information Requirements (CCIR)
	Common Map Display
	Picture Consistency
	Operations Overlay Feedback
Shared Situational Awareness	Map Area
	Sensor Coverage
	Satellite Coverage
	Line of Sight
	Surprise Attack
	Collateral Damage
	SITREP Lag
Communication	Whiteboard Use
	Radio Communications Pattern
	Personnel Initiating Whiteboard Conferences

(table continues)

Table 5 (Continued)

Team Process Skill Dimension	Candidate Measure of Performance
Coordination	Overlay Use
	Whiteboard Commonality
	Targeting
	Fire Support Coordination
	Fire Engagements
	Opposing Forces (OPFOR) Destruction
Decision-Making	Unmanned Aerial Vehicle (UAV) Effectiveness
	Length of Battalion Decision-Making Cycle (Operations Order)
	Length of Battalion Decision-Making Cycle (Platoon Movement)

Nine of the candidate measures were similar to and/or adapted from those implemented during the BCR III (see Table 3, Measures 2a, 2d, 2e, 3b, 4a, 4b, 4c, 5a, and 5b). Others were similar to and/or adapted from measures developed by Dzindolet et al. (1998) and Mason (1995). The rest of the measures were developed based on the lessons learned from BCR III discussed in Throne et al. (1999). In total, there were 28 candidate measures selected for development, as listed in Table 5. For each of these measures, draft operational definitions, rationales, and output formats were developed. These are all provided in Appendix B.

In an attempt to provide more useful feedback to the staff, output formats for the measures were designed to be as simple yet informative as possible. Based on the lessons learned during BCR III (Throne et al., 1999), the output formats for each of the measures were also specified in order to allow the programmers to better meet the Project Team's expectations of providing clearer feedback to the staff on their performance. All performance measures were designed so their outputs could be displayed in at least one of these three formats: tables, graphs, and pictures.

An emphasis was placed on designing pictorial outputs since meaningful feedback can be provided to users through the use of automated pictorial comparisons (Lickteig & Throne, 1999). With instrumented C^4I systems, automated pictures that capture and compare users' tactical screens can potentially be developed. As an example, Lickteig and Throne depict a notional situation where the C^4I display of a company commander shows four enemy tanks where the battalion commander's C^4I display does not. Ideally, this discrepancy between the two levels of command would be automatically detected by the system and presented to the unit in picture format immediately following the training exercise as a form of extrinsic performance feedback. Other notable examples of pictorial formats for unit and staff performance are provided by Brown et al. (1998). The Project Team designed a series of measures to capitalize on this capability of providing pictorial measure formats that directly relate performance to the tactical context in which the performance occurred (e.g., Map Area and UAV Effectiveness).

The ability of the SC^4 system to provide intrinsic training feedback on staff performance was also considered. Intrinsic training feedback is feedback provided during an exercise and is obtained by directly observing actual and simulated vehicles and units and their associated activities (Brown et al., 1998). The SC^4 system, however, was not yet designed to provide intrinsic feedback on team process skills directly, which was the focus of this research. Therefore, the Project Team did not consider intrinsic feedback as an option at this stage of research in designing automated measures. All the automated measures designed for this project are extrinsic measures. That is, they provide feedback after an exercise has ended, similar to an AAR.

Once the candidate measures were designed, they were passed to the SAG for review and approval. For each candidate measure, the SAG was given the information contained in Appendix B: measure name, operational definition, rationale, recommended output formats, and the team process skill dimension that could be partially measured. The SAG provided a consensus rating for each candidate measure. The consensus rating indicated whether the measure should be developed now, developed later, or developed last. The SAG determined that all but one measure should be developed now.

In addition, the SAG was also asked to provide any additional measures they felt needed to be added to the list. After they met and discussed the candidate measures, two additional measures were added to the list. The first measure of performance was named *Information Retrieval by the Commander*, and fell under the team process skill dimension of Decision-Making. The second measure of performance was named *Information Flow*, and fell under the team process skill dimension of Communication.

Based on the SAG's ratings, the Project Team met with MMBL system programmers to begin developing the measures. After meeting with the programmers, some of the measures were considered to be too time-consuming or expensive to develop, given the BCR IV time and cost constraints. Of the 28 candidate measures and two additional SAG measures originally proposed for development, 19 measures were selected as feasible for development, given the BCR-related costs and time constraints. These measures are described next.

Development

After reaching a consensus on the 19 measures to develop, the MMBL system programmers were given the list of measures summarized in Table 6, including an operational definition, rationale, and recommended output formats for each measure (see Appendix B).

Table 6

Automated Measures Developed for Battle Command Reengineering IV by Skill Dimensions

Skill Dimension and Measure Name	Description
Adaptability	
Terrain Analysis	Amount of time each duty position uses each of the following tools: Stealth Control, Terrain Intervisibility, FOV, Snail Display, and FLOT Display during the mission.
Performance Monitoring and Feedback	
SITREP Use	Frequency and duration of use of the SITREP tool during each mission by duty position.
SPOTREP Use	Frequency and duration of use of the SPOTREP tool during each mission by duty position.
CCIR	Frequency and duration of use of the CCIR tool during each mission by duty position.
Shared Situational Awareness	
Map Area	Square kilometers of the battlefield displayed and center point of each resized PVD map during each mission by duty position at critical points in time.
Line of Sight	Frequency and duration of use of the Line of Sight tool during each mission by duty position.
Surprise Attack	Total number of flank or rear engagements on OPFOR and BLUFOR vehicles during each mission.
Collateral Damage	Total number of attacks on BLUFOR non-instrumented vehicles and/or personnel by indirect non-line of sight weapons systems under battalion control during each mission.
Communication	
Whiteboard Use	Total number of Whiteboard files residing on each workstation for each mission by duty position.
Radio Communications Pattern	Frequency and duration of use of battalion command and operations-intelligence radio nets and Whiteboard conferencing during each mission by duty position at critical points in time.
Personnel Initiating Whiteboard Conferences	Total number of Whiteboard conferences, lasting 3 minutes or more, initiated during each mission by duty position.

(table continues)

Table 6 (Continued)

Skill Dimension and Measure Name	Description
Coordination	
Overlay Use	Total number of workstations showing the same operations overlay file that the squadron commander is showing on his PVD for each mission by duty position at critical points in time.
Whiteboard Commonality	Total number of Whiteboard directories showing the same Whiteboard files that the commander and deputy commander have for each mission by duty position at critical points in time.
Targeting	Total number of times a SPOTREP query was conducted on the target identified in a fire support request immediately before it was transmitted by a staff member.
Fire Support Coordination	Ratio of OPFOR kills due to indirect fire from units controlled by the squadron staff to OPFOR kills due to direct fire controlled by squadron subordinate units.
Fire Engagements	Average range of OPFOR and BLUFOR kills by vehicle type during each mission.
OPFOR Destruction	Time from the first OPFOR engagement until OPFOR vehicle losses exceeded 70%. In addition, the rate at which the OPFOR was killed in five minute intervals from the first engagement until the last OPFOR kill during each mission.
Decision-Making	
UAV Effectiveness	Percentage of OPFOR vehicles first detected by the UAV under squadron control for each UAV launch during each mission.
Information Retrieval by the Commander	Type and frequency of information the commander retrieves on his own that other staff normally retrieve for him.

Note. BLUFOR = blue forces; CCIR = commander's critical information requirements; FLOT = forward line of troops; FOV = field of view; OPFOR = opposing forces; PVD = plan view display; SITREP = situation report; SPOTREP = spot report; UAV = unmanned aerial vehicle.

Unfortunately, pictorial and graphical representations for most of the automated measures could not be supported by the MMBL's programming analysts. Therefore, most of the automated measures outputs from the DCA were developed in table form. However, the Project Team attempted to develop prototype pictorial formats based on the data tables received from the DCA. For the measures that were developed with a pictorial format, production of the outputs was quite time-consuming. These picture formats were all developed in a commercial off-the-shelf spreadsheet software program, Microsoft® Excel, by the Project Team. Iterative experimentation with the picture formats was required to relate the DCA output data to the operational context. Finally, since these formats were not created in the DCA, the actual terrain and vehicle icons could not be represented in the pictures. Therefore, any vehicles depicted in the pictures presented in the Battle Command Reengineering IV Results and Discussion section are represented by symbols, and no terrain features appear in the pictures.

In addition, an automated flag list was developed as a tool to assist in determining critical points during BCR IV missions for which data needed to be collected. This flag list was based on the CVCC (Leibrecht et al., 1994) and TARGETs (Fowlkes et al., 1994) methodologies. The intent was to identify critical events around which the staff could be expected to be the most proactive in supporting the commander's decision-making process and which could be automatically detected by the DCA. By using an automated flag list, the Project Team hoped to eliminate the need for observers to record the times of critical events. The automated flags would also allow extraneous trial data to be filtered out by the DCA programmers prior to the start of the analysis process. This list was given to the programmers prior to the start of BCR IV. Table 7 provides a list of the automated flags identified and developed.

Table 7

Automated Flags for the Data Collection and Analysis System

1. Start of exercise	9. Last engagement
2. End of exercise	10. C^2 node loss
3. First detection of OPFOR	11. PVD operations overlay file opened
4. First indirect fire engagement	12. NLOS engagement of non-instrumented systems
5. First direct fire engagement	13. Whiteboard conference exceeding 3 minutes
6. First friendly casualty	14. Friendly platoon status change (green, amber, red, black)
7. Friendly losses exceed 30%	
8. OPFOR losses exceed 70%	15. Friendly platoon location change greater than 2000 meters

Note. C^2 = command and control; NLOS = non-line of sight; OPFOR = opposing forces; PVD = plan view display.

The Project Team also developed a manual exercise flag log, based on BCR III experience as well as for redundancy in case the automated flags failed to operate properly. The manual log had three functions: a) to serve as a means to verify that the automated flags could in fact automatically identify the timing of critical events; b) to identify the key events and times if automated event flags could not be derived; and c) to provide a means to record significant events which had not been anticipated to occur during a trial. The manual exercise flag log used for BCR IV can be seen in Figure 5.

The Project Team provided one observer during BCR IV. Four other observers from the Test and Evaluation Coordination Office volunteered to maintain an exercise flag log as their primary duties allowed. The OPFOR cell also maintained an exercise flag log. Both groups used the log shown in Figure 5. Each observer received a short briefing prior to the start of the experiment from the Project Team on the need for the exercise flag log and how to complete it.

EXERCISE FLAG LOG

Today's Date _____ **Exercise ID** _____

BATTALION AREA OF INTEREST/RESPONSIBILITY

Coordinates of upper left hand corner(_____)
Coordinates of lower left hand corner(_____)
Coordinates of upper right hand corner(_____)
Coordinates of lower right hand corner(_____)

BATTALION AREA OF OPERATION

Coordinates of upper left hand corner(_____)
Coordinates of lower left hand corner(_____)
Coordinates of upper right hand corner(_____)
Coordinates of lower right hand corner(_____)

Start of Mission (Time [HH:MM:SS] ____:____:____)

Receipt of Brigade order (Time of Receipt [HH:MM:SS] ____:____:____)

Start/end of planning
(Start time of mission planning [HH:MM:SS] ____:____:____)
(Stop time of mission planning [HH:MM:SS] ____:____:____)

Start of mission rehearsal (Time [HH:MM:SS] ____:____:____)

Start/end of execution phase
(Start time of mission execution [HH:MM:SS] ____:____:____)
(Stop time of mission execution [HH:MM:SS] ____:____:____)

Cross line of departure (Time [HH:MM:SS] ____:____:____)

First detection of OPFOR (Time [HH:MM:SS] ____:____:____)

First direct fire engagement (Time [HH:MM:SS] ____:____:____)

First indirect fire engagement (Time [HH:MM:SS] ____:____:____)

Friendly losses exceed 30% (Time [HH:MM:SS] ____:____:____)

Enemy losses exceed 70% (Time [HH:MM:SS] ____:____:____)

Last engagement (Time [HH:MM:SS] ____:____:____)

C^2 node combat loss (Time [HH:MM:SS] ____:____:____)

Report of completion
(Time mission completion reported [HH:MM:SS] ____:____:____)

Start of staff consolidation
(Start time of staff consolidation [HH:MM:SS] ____:____:____)

End of mission (Time [HH:MM:SS] ____:____:____)

Mission Success (White Cell Estimate Yes ____/No____)

Total Number of Vehicles

Total Number of OPFOR _____

Tanks _____

IFV _____

APC _____

ARTY _____

Total Number of BLUFOR _____

System Crashes *System(s)*_____

Time system went down (HH:MM:SS) ____:____:____

Time system back up (HH:MM:SS) ____:____:____

OTHER SIGNIFICANT EVENTS

Event _____ Time (HH:MM:SS) ____:____:____

Event _____ Time (HH:MM:SS) ____:____:____

Event _____ Time (HH:MM:SS) ____:____:____

Figure 5. Exercise flag log for Battle Command Reengineering IV.

Measures Implementation Process

The implementation of the automated measures developed to partially assess team process skill dimensions was dependent upon the BCR IV environment. Therefore, descriptions of the BCR IV participants, BCR missions, and the implementation of the measures in the BCR IV experiment follow.

Participants

The unit participating in the experiment was an active Army cavalry squadron staff with its subordinate company commanders participating. One company brought drivers and gunners to man several Future Combat Vehicle simulators. Table 8 shows the call signs and associated node positions for the commander and 13 primary staff members. The battle command reengineering aspect of the BCR IV experiment was focused on this group of 14 soldiers. In this report, these participants will often be identified by their call signs, especially in tables that appear in the Battle Command Reengineering IV Results and Discussion section. The node functions and job responsibilities for each staff member were left to the discretion of the squadron commander, who was allowed to reorganize the staff as he gained experience in operating the SC^4 system.

Table 8

Staff Member Call Signs and Titles

Call Sign	Title	Node
WP6	Squadron Commander	Command 1
WP62	Enemy Operations Officer	Command 1
WP63	Effects Officer	Command 1
WP5	Deputy Squadron Commander	Command 2
WP52	Operations NCO	Command 2
WP53	Operations Officer	Command 2
WP88	Battle Captain	Control 1
WP82	Enemy Operations NCO	Control 1
WP83	Friendly Operations Officer	Control 1
WP84	Sensor NCO	Control 1
WP99	Battle Captain	Control 2
WP92	Enemy Operations NCO	Control 2
WP93	Friendly Operations Officer	Control 2
WP94	Sensor NCO	Control 2
Iron 6	Company Commander	Constructive Simulation
Killer 6	Company Commander	Constructive Simulation
Lightning 6	Company Commander	Constructive Simulation
Mad Dog 6	Company Commander	Constructive Simulation

Note. NCO = non-commissioned officer.

Other squadron personnel involved in the experiment included: six company commanders, six deputy company commanders, three maneuver platoon leaders, one mortar platoon leader, one scout platoon leader, one scout platoon sergeant, five gunners, and eight drivers.

Battle Command Reengineering Missions

The BCR IV experiment missions were based on tactical operations that an Army battalion, equipped with an advanced digital C^4I system, might be expected to conduct in the year 2010 and beyond. The virtual terrain chosen for the experiment was northeastern Bosnia-Herzegovina, centered around the city of Tuzla. This terrain is extremely mountainous with limited ground mobility corridors. Figure 6 shows the experiment terrain map with a battalion area of operations superimposed.

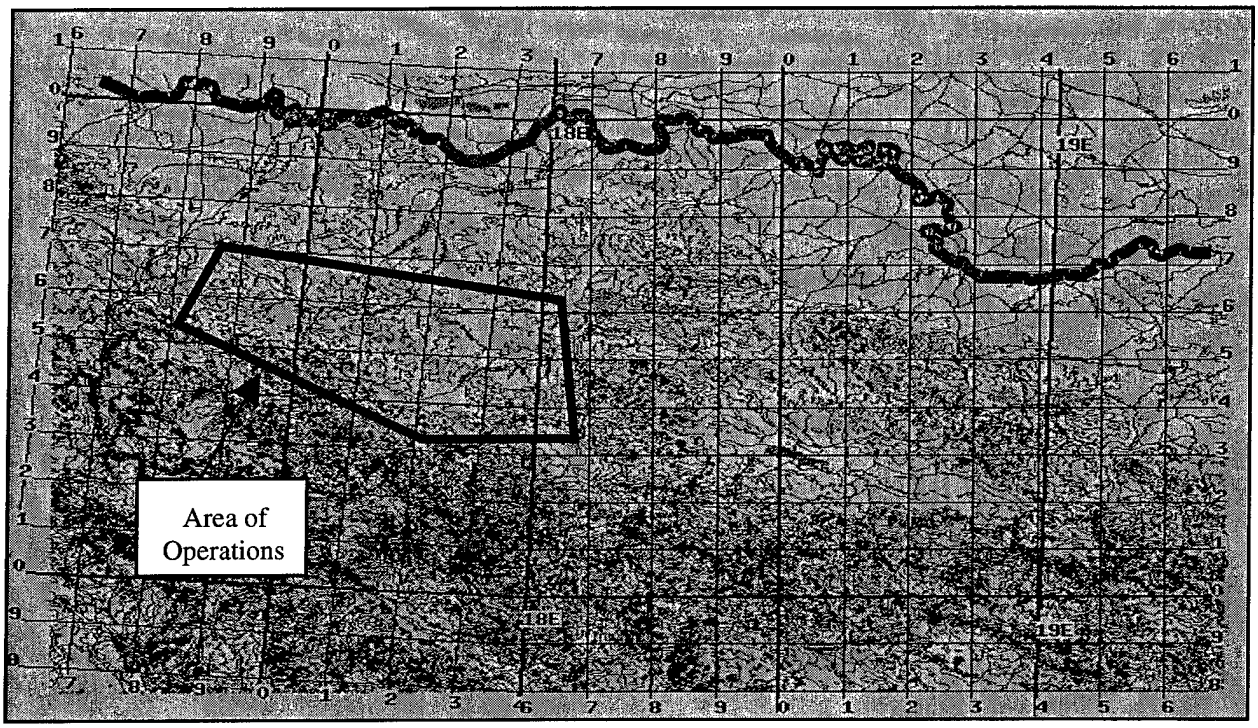


Figure 6. Battle Command Reengineering IV terrain map.

Implementation

The automated measures were not implemented during conduct of the BCR IV, but rather were to be applied to the DCA recordings of the exercise. After completion of the BCR IV, MMBL programmers worked on providing the automated measures data on the team process skill dimensions. However, due to the volume of data, additional data collection formats had to be designed in order to pull the data needed from the system. After this initial filtering and data reduction, sample data were passed on to the Project Team for review. From that point forward, MMBL programmers and members of the Project Team worked together to develop data sets that were satisfactory for analysis. When the data had been filtered to an appropriate level, MMBL programmers passed the complete data files to the Project Team for subsequent analysis.

As expected, after receiving the data files from the MMBL programmers, more off-line filtering of the DCA data was required to extract the automated measures of performance (shown in Table 6). For example, preliminary radio system analyses indicated numerous radio transmissions of 1 second or less in length. Since it was determined by the Project Team that these data contained no useful information, the data were filtered again by the MMBL programmers to exclude these transmissions from the analyses.

Additionally, some of the data, as indicated in Table 6, were filtered to only examine specific points in time that were considered to be critical. Critical times were derived from the automated flags presented in Table 7, such as time of first indirect fire engagement or time of first friendly casualty. When the MMBL programmers and SMEs from the Project Team attempted to derive the critical times for the automated flags, it was determined that it would be very time-consuming to design a program to accomplish this. Therefore, an MMBL programmer

and a SME from the Project Team worked together and went through the DCA master event data log for each mission, found the desired critical events, and recorded the times manually. When these times from the DCA were compared to the times manually recorded by the observers, there appeared to be no discrepancies.

In summary, a prototype automated measurement package of team process skills was developed. However, true implementation of automated measures as defined by the Project Team was not achieved. Notably, considerable progress was made in matching DCA output data to team process skill dimensions. Additionally, a start was made on converting measures data into pictorial representations of staff performance. Due to a variety of technical and tactical reasons, not all measures designed by the Project Team were developed and implemented during BCR IV. A discussion of the results obtained from the prototype measures package implemented during BCR IV follows.

Battle Command Reengineering IV Results and Discussion

Sample results obtained from the automated measures will be presented in the framework of Cannon-Bowers et al.'s (1995) team process skill dimensions. These samples provide prototype examples for some of the automated measures developed. Additional results are provided in Appendix B. A more comprehensive set of results was provided to the MMBL for inclusion in their final report (MMBL, in preparation). The focus of this section is to document and examine the prototype automated measures developed.

The Project Team made no attempt to comprehensively assess whether the squadron had accomplished its assigned tasks during each mission discussed in this report. While team processes need to be linked to outcomes for a comprehensive understanding of team performance (Cannon-Bowers & Salas, 1997; Smith-Jentsch et al., 1998), assessing the overall squadron effectiveness during tactical missions was beyond the scope of this project. In fact, the BCR IV experimentation focus during the missions reported was to gain insights into streamlined battle command with a future force across doctrinal, training, leadership, and soldier issues (MMBL, in preparation), not unit performance. An overall performance assessment of the unit was not made by the MMBL for each mission nor for the BCR IV experiment as a whole.

Data supporting the automated measures were gathered by the MMBL's DCA during BCR IV's six trial missions. After considering the available data, the MMBL and Project Team limited the primary data analysis to four of the six missions, as indicated in Table 9. Mission 1 data were excluded because of the unique combination of tactical operations that were not repeated during the rest of the experiment. Mission 6 was excluded because there were only two hours of data available for analysis. The data from the analyzed missions were provided in a spreadsheet format by DCA programmers. They were then reduced, compiled, and converted into tables, graphs, and/or pictures by the Project Team. After the output formats were finished, a quality control check was conducted to make sure all calculations were correct and the outputs contained the intended data to be reported.

Table 9

Battle Command Reengineering IV Experiment Missions

Mission	Tactical Description	Remarks
1[a]	Rearward passage of lines, relief in place, defend against dismounted attack	Multiple operations not repeated during experiment
2	Defend in sector	
3	Move to engage	
4	Defend in sector	
5	Move to engage	
6[a]	Move to engage	Two hours of data

[a]Missions excluded from analysis by Mounted Maneuver Battlespace Lab and Project Team.

The main purpose of this project was to develop prototype automated measures of staff processes, not to rate the performance of the staff involved in BCR IV. Therefore, this section will not focus on the staff's performance. Rather, it will summarize the utility of the prototype measures designed to partially assess staff processes, and the attempt to depict those processes in more meaningful formats. It should be noted that the measures not developed are potentially doable in future efforts. However, due to the BCR-related costs and time constraints, they were not feasible for this project. Finally, all results presented in this report are exploratory and formative.

Adaptability

Adaptability is "the process by which a team is able to use information gathered from the task environment to adjust strategies through the use of compensatory behavior and reallocation of intrateam resources," as indicated in Table 1. Adaptability may also be referred to as flexibility, compensatory behavior, or dynamic reallocation of functions (Cannon-Bowers et al., 1995). The three candidate measures defined and presented to the SAG as ways to partially assess adaptability included Terrain Analysis, Node Location, and Loss of Node. The Project Team's premise was that by looking at how the staff obtained and shared terrain data, where they located themselves physically in relation to their subordinate units, and what actions they took in reaction to a combat loss of some of the staff, may indicate their adaptability to unforeseen or changing circumstances.

Of the three candidate measures, only the Terrain Analysis measure was developed. All conclusions regarding measurement of Adaptability are based on this measure alone.

Terrain Analysis

Terrain Analysis was measured in terms of the amount of time each staff member used certain PVD tools, including Field of View (FOV), Snail Display, Stealth Control, Terrain Intervisibility, and Forward Line of Troops (FLOT) for each mission. The goal was to examine whether these tools were used to obtain an appreciation of the effects of the terrain on mission planning, rehearsal, and execution. Two output formats were designed for this measure: graphs,

which displayed the average time of use of each tool for each mission; and tables, which for each tool displayed the frequency and duration of use by each staff member for each mission. The FLOT tool was never used by any of the staff members during any of the missions. Figures 7 through 10 show the average number of seconds for each time the tools were used by the primary staff members for each mission.

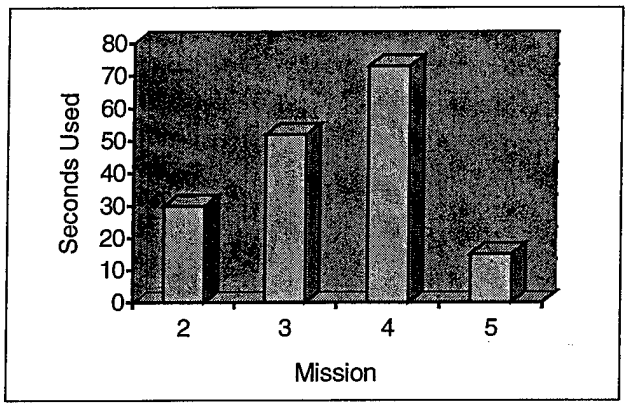


Figure 7. Average duration of each use of the Field of View tool, by mission.

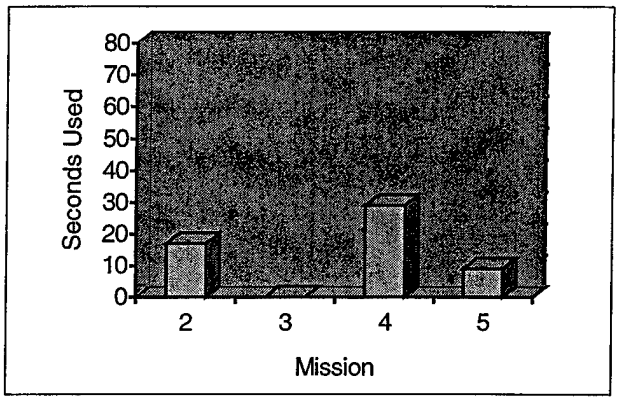


Figure 8. Average duration of each use of the Snail tool, by mission.

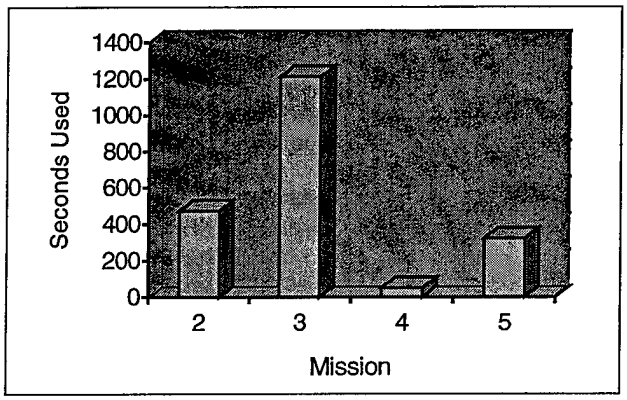


Figure 9. Average duration of each use of the Stealth tool, by mission.

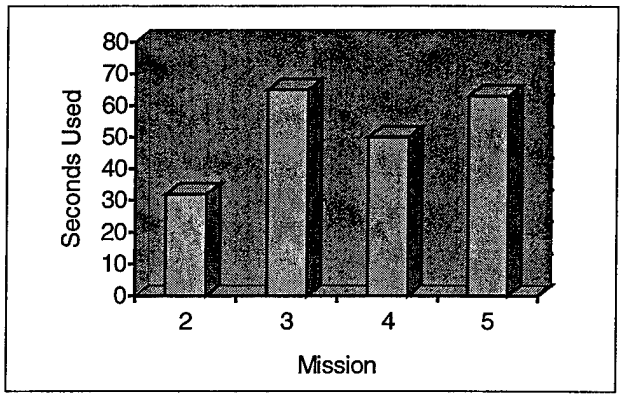


Figure 10. Average duration of each use of the Terrain Intervisibility tool, by mission.

In Figure 7, it can be seen that the FOV tool was used for increasing periods of time from the second mission to the fourth mission, then went down again for the fifth mission. It was used by a total of six of the primary staff members (see Appendix B). The Snail tool was used for much shorter periods (see Figure 8) and it was used by only three of the primary staff members. The Stealth tool was used for greater durations of time than the other tools (see Figure 9), but its usage was limited to five of the primary staff members. Average duration of use for the Terrain Intervisibility tool almost doubled after Mission 2 (see Figure 10). Six of the primary staff members used this tool.

A sample of the tables produced for this measure appears in Table 10, which shows the frequency and duration of use of the Terrain Intervisibility tool for each staff member who used the tool by mission. As seen in Table 10, six of the primary staff members used the tool, ranging anywhere from 22 seconds to 14 minutes, 34 seconds. More detailed information on the commander's and each primary staff member's use of these tools can be seen in Appendix B.

Table 10

Frequency and Duration of the Terrain Intervisibility Tool Use for Staff Members by Mission

Position	Mission 2		Mission 3		Mission 4		Mission 5	
	Frequency	Duration	Frequency	Duration	Frequency	Duration	Frequency	Duration
WP62	0	0:00	1	5:27	1	3:21	0	0:00
WP52	9	4:31	2	0:29	0	0:00	13	14:34
WP53	3	0:39	4	2:03	10	1:25	5	1:12
WP82	1	1:19	2	0:22	2	1:40	1	2:03
WP83	4	2:38	3	4:41	6	4:59	2	2:42
WP92	0	0:00	0	0:00	7	10:10	2	3:42

Note. Duration times are total use in minutes and seconds.

Summary of Adaptability

Because of the limited amount of data available for analysis, the Project Team cannot make an overall assessment of utility for the Adaptability measure implemented. Nevertheless, the Terrain Analysis measure did provide some useful information, especially to SC^4 system designers and trainers. This information was passed to the MMBL for their review. Based on the results obtained during BCR IV, the Project Team concluded that the Terrain Intervisibility tool use measure, which was designed to measure to behaviors related to back-up plans and quick adjustment to situational change, might be more relevant to the shared Situational Awareness dimension and its associated behaviors. The Node Location and Loss of Node measures may have provided insights into the staff's adaptability if they had been implemented.

Performance Monitoring and Feedback

Performance Monitoring and Feedback is "the ability of team members to give, seek, and receive task-clarifying feedback," as indicated in Table 1. It may also be referred to as

intramember feedback or mutual performance monitoring (Cannon-Bowers et al., 1995). The six candidate measures defined and presented to the SAG as ways to partially assess Performance Monitoring and Feedback included Situation Report (SITREP) Use, Spot Report (SPOTREP) Use, Commander's Critical Information Requirements (CCIR)[4], Common Map Display, Picture Consistency, and Operations Overlay Feedback. The Project Team's premise was that the use of these tools might allow the staff to keep track of each others' and the unit's activities. If these tools are not used, it could be inferred that the staff is not monitoring performance and therefore, cannot provide adequate feedback to each other.

Of these six candidate measures, only SITREP Use, SPOTREP Use, and Priority Information Request (PIR) were developed. All conclusions regarding Performance Monitoring and Feedback are based on these three measures. The SPOTREP data are provided as an example of the output received from the Performance Monitoring and Feedback measures implemented during BCR IV. Data from the other measures are provided in Appendix B.

Spot Report Use

The SPOTREP Use measure was designed to determine how often the staff monitored OPFOR units' activities and statuses throughout the missions. The SPOTREP information usually includes the size, activity, location, time, and equipment of the OPFOR being reported on. Table 11 shows the frequency and duration for use of the SPOTREP tool during the BCR for each staff member who used the tool. Durations are presented in minutes and seconds and represent the total duration the tool was used for each mission. From the table, it appears that the SPOTREP tool was used more frequently as time progressed. This is also indicated in Figure 11 and Figure 12, which show the average frequency of use across the commander and 13 primary staff members and by node, respectively.

Table 11

Frequency and Duration of Spot Report Use for Staff Members by Mission

Position	Mission 2		Mission 3		Mission 4		Mission 5	
	Frequency	Duration	Frequency	Duration	Frequency	Duration	Frequency	Duration
WP63	8	00:20	8	00:25	9	00:29	0	00:00
WP52	0	00:00	18	00:34	23	01:54	0	00:00
WP53	0	00:00	12	00:19	21	00:37	23	02:47
WP88	0	00:00	1	00:03	0	00:00	19	03:49
WP82	0	00:00	21	14:05	14	05:10	0	00:00
WP83	5	00:19	0	00:00	28	48:04	22	04:06
WP99	3	00:06	0	00:00	0	00:00	3	01:13
WP92	15	02:09	0	00:00	0	00:00	0	00:00
WP93	5	00:19	0	00:00	46	14:48	12	01:18

Note. Duration times are total use in minutes and seconds.

[4] The CCIR tool was renamed the Priority Information Request (PIR) tool prior to the start of BCR IV. Future references to this tool will refer to the PIR tool.

As seen in the figures, the SPOTREP was used most often during Mission 4. Its frequency of use progressed from one mission to the next, peaked at Mission 4, and then declined during Mission 5. Command 2 (WP52 and WP53) seemed to use the SPOTREP more often than the other nodes. However, as seen in Table 11, Control 1 (WP88, WP82, and WP83) used the SPOTREP for the longest periods of time.

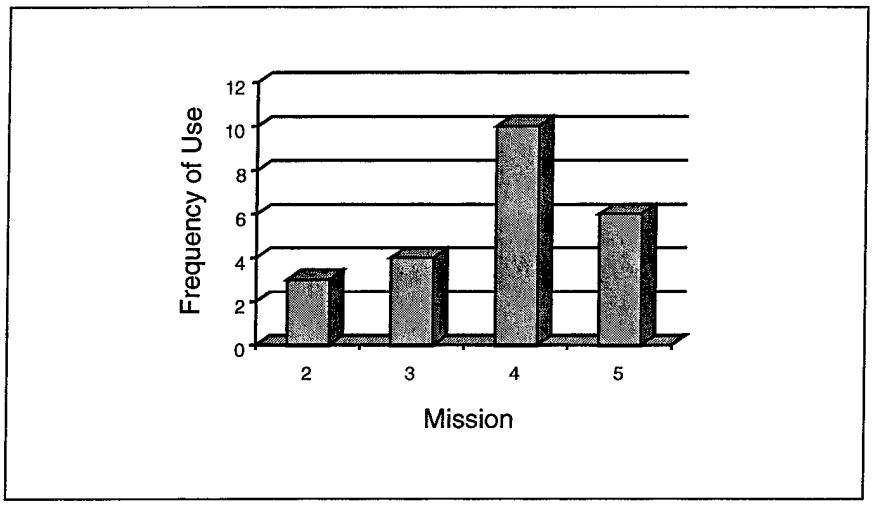


Figure 11. Average frequency of spot report use for commander and staff by mission.

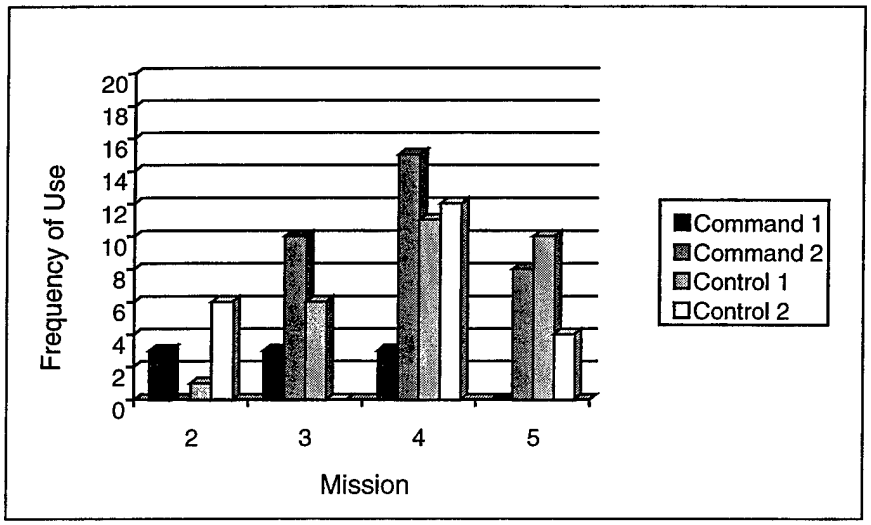


Figure 12. Average frequency of spot report use for each node by mission.

Summary of Performance Monitoring and Feedback

The combination of proposed automated measures utilized during this project (SITREP Use, SPOTREP Use, and PIR) to provide insight into the skill dimension of Performance Monitoring and Feedback appeared useful. Between the capability of the SC^4 system to automatically display information about friendly and OPFOR units and the widespread usage of the SITREP, SPOTREP, and PIR tools across all four command and control nodes, the staff had and used the

tools they needed to monitor the performance of the squadron and provide recommendations to the commander and guidance to subordinate company commanders.

The data from these measures did provide some insight into the behaviors that are associated with this skill dimension. The actions that staff members were taking to obtain information and to track the performance of other staff members could be measured. The inclusion of an additional automated measure such as the Radio Communications Pattern measure described under the Communication skill dimension in Appendix B would provide a fuller picture of monitoring and feedback activities within the staff. Implementation of the Picture Consistency and Operations Overlay Feedback measures, which will require extensive programming to develop, should be a priority.

Shared Situational Awareness

Shared Situational Awareness is "the process by which team members develop compatible models of the team's internal and external environment," as indicated in Table 1. This may also be referred to as situational awareness or shared problem-model development (Cannon-Bowers et al., 1995). The seven candidate measures defined and presented to the SAG as ways to partially assess shared situational awareness included Map Area, Sensor Coverage, Satellite Coverage, Line of Sight, Surprise Attack, Collateral Damage, and SITREP Lag. For a majority of the candidate measures for this skill dimension, pictorial output formats were recommended for development. The Project Team's premise was that the results from these measures might indicate how aware staff members were overall of what was going on during the missions.

Of these seven candidate measures, only Map Area, Line of Sight, Surprise Attack, and Collateral Damage were developed. All conclusions regarding Shared Situational Awareness are based on these four measures. Only one of the candidate pictorial formats was developed: the Map Area measure. The Map Area data are provided as an example of the output received from the Shared Situational Awareness measures implemented during BCR IV. Data from the other measures are presented in Appendix B.

Map Area

The Map Area measure was designed to determine where on the battlefield the commander and primary staff members were focused at critical points in time and how much of the battlefield they were viewing on their PVDs. Map Area results were captured in two ways: a table with the average amount of battlefield each staff member displayed for each mission (see Table 12), and pictorial representations of the center point of each staff member's PVD display at critical points in time for each mission (see Figure 13).

Table 12

Average Km^2 of Battlefield Displayed by Each Staff Member for Each Mission

Position	Mission 2	Mission 3	Mission 4	Mission 5
WP6	10,146	8,214	12,504	15,201
WP62	1,402	752	1,507	6,285
WP63	988	853	1,279	5,645
WP5	5,515	2,876	3,838	5,809
WP52	1,736	3,519	2,245	3,410
WP53	1,030	781	1,419	989
WP88	5,059	3,986	2,770	5,090
WP82	1,934	626	1,872	3,327
WP83	5,173	3,499	5,122	5,644
WP99	2,056	1,058	2,702	2,957
WP92	3,846	6,858	6,040	6,274
WP93	1,146	497	1,730	1,958
Squadron Area of Operations	2,260	2,640	2,260	5,200
Squadron Area of Interest	3,070	4,200	3,070	8,925

Table 12 shows the average amount (in km^2) of battlefield displayed by each staff member for each mission. As indicated in the table, the squadron commander (WP6) had more of the battlefield displayed on his screen than the other staff members. The amount of battlefield he was looking at overall was usually 2-3 times the squadron area of operations or even the squadron area of interest. The rest of the staff members' displays were generally focused on much smaller areas of the battlefield than the commander. However, during Mission 5, the majority were looking at a larger portion of the battlefield than during the earlier missions. This may be explained by the increase in the squadron's area of responsibility in Mission 5.

Figure 13 is a representation of the entire area of the battlefield for one mission (Mission 2) at a sample point in time (first indirect fire). The center points for each staff member's PVD display at the moment of first indirect fire are marked. The larger box represents the squadron commander's view, and the smaller box represents a company commander's field of view. Finally, the squadron's area of operations is also marked.

This figure represents an initial attempt at presenting data in a pictorial format. Perhaps more detailed and interrelated information can be gained from such a picture than from a subjective description of the event by observers or tabular formats, such as Table 12, especially for feedback to the staff. For example, it is evident from the figure that most of the staff members had the center points of their maps set in nearly the same location. In addition, it appears that, at this point in the fight, the squadron commander was as interested in seeing where the OPFOR was as he was in viewing his own unit.

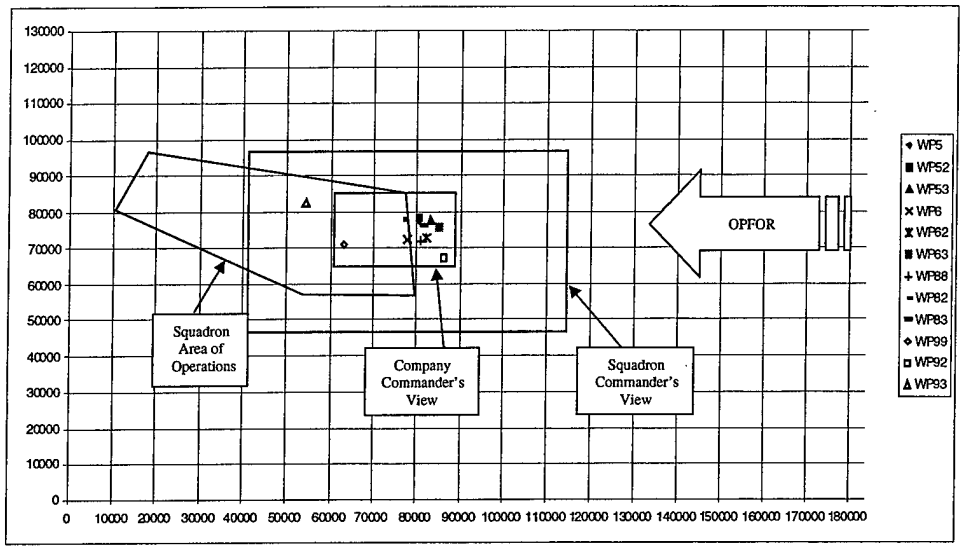


Figure 13. Center points for each staff member's Plan View Display at first indirect fire for Mission 2, with the squadron commander's and a company commander's views.

Similar pictures for other critical points (start of exercise [STARTEX], first direct fire, and end of exercise [ENDEX]) during Mission 2 can be seen in Figures 14-16. By comparing all four figures, it appears that the center points of staff members' screens were dispersed at the beginning of the mission, but would become centered in the same area once a critical event (e.g., first indirect or direct fire) occurred. It is interesting to note that by the end of the mission, the center points of everyone's field of view had shifted considerably in the direction from which the OPFOR was coming. This may indicate that the staff felt that the immediate battle was under control and they were trying to focus on new threats not yet identified.

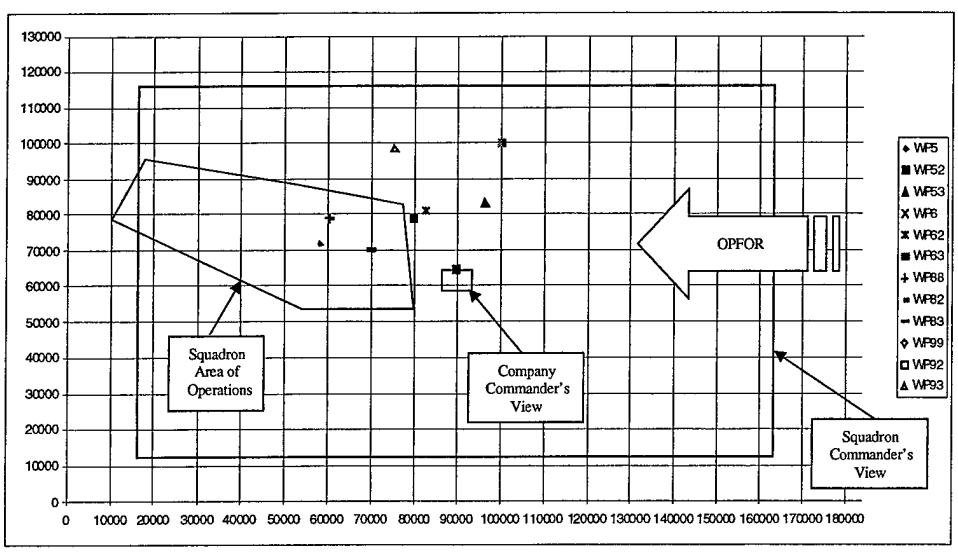


Figure 14. Center points for each staff member's Plan View Display at start of exercise for Mission 2, with the squadron commander's and a company commander's views.

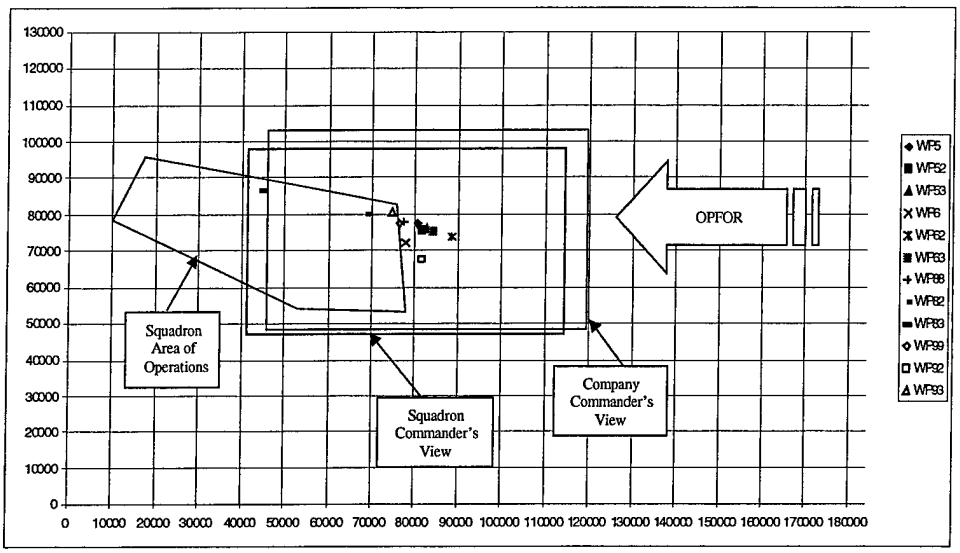


Figure 15. Center points for each staff member's Plan View Display at first direct fire for Mission 2, with the squadron commander's and a company commander's views.

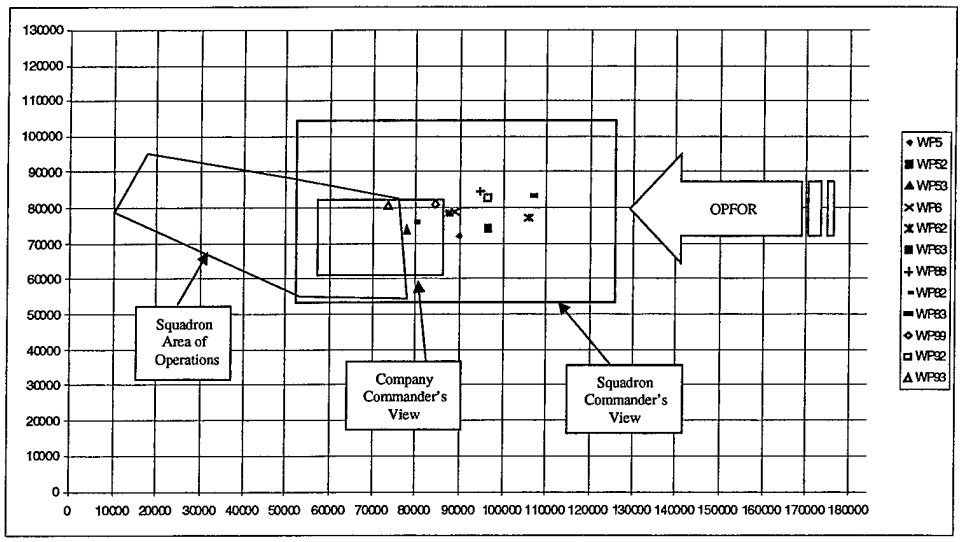


Figure 16. Center points for each staff member's Plan View Display at end of exercise for Mission 2, with the squadron commander's and a company commander's views.

Summary of Shared Situational Awareness

A partial understanding of whether the staff had shared situational awareness could be developed from review of the data on the four measures (Map Area, Line of Sight, Surprise Attack, and Collateral Damage). The pictorial format associated with the Map Area measure provided more information than the associated table. The additional data from the measures not implemented during BCR IV, especially SITREP Lag, might have added more support to an assessment of the staff's shared situational awareness.

Communication

Communication is "the process by which information is clearly and accurately exchanged between two or more team members in the prescribed manner and with proper terminology," as indicated in Table 1. This may also be referred to as information exchange or consulting with others (Cannon-Bowers et al., 1995). The three candidate measures defined and presented to the SAG as ways to partially assess Communication included Whiteboard Use, Radio Communications Pattern, and Personnel Initiating Whiteboard Conferences. The Project Team's premise was that the results from these measures might suggest when and how often staff members communicated with each other.

A fourth measure was added by the SAG: Information Flow. Of these four candidate measures, only Whiteboard Use, Radio Communications Pattern, and Number of Staff Personnel Initiating Whiteboard Conferences were developed. All conclusions regarding Communication are based on these three measures. Data from the Whiteboard Use measure are provided as an example. Additional data collected on the Radio Communications Pattern and Initiating Whiteboard Conference measures can be found in Appendix B.

Whiteboard Use

The SC^4 Whiteboard tool is a system used by two or more staff officers in the same or in separate nodes to conduct simultaneous collaborative planning while looking at the same tactical picture. Each staff member can draw on or modify the picture, and other staff members will be able to view that drawing or writing as it is done. This system was also used to distribute written squadron operations orders (OPORDs) during BCR IV. The only way that OPORDs could be saved on an individual SC^4 system was if the commander or staff member performed the correct sequence of operations to save the Whiteboard file containing the OPORDs. The number of saved Whiteboard files on a particular SC^4 system indicates the extent to which the commander or staff member thought the Whiteboard file was sufficiently important that it should be kept for reference after a Whiteboard conference was completed. Therefore, the intent of this measure was used to assess the extent to which the staff was using the Whiteboard tool, as an indicator of its efforts to coordinate its activities.

Table 13 shows how many Whiteboard files each staff member had saved on the SC^4 system for each mission. As shown in the table, almost everyone except the squadron commander (WP6) and the sensor NCOs (WP84 and WP94) had at least one Whiteboard file residing on his SC^4 system at the end of each mission. A notable finding is that the squadron commander did not have any Whiteboard files on his SC^4 system. The data does not provide a ready explanation for this. Another finding was that the Whiteboards for the missions were more likely to be distributed to the company commanders on the later missions than on the earlier ones. Staff members in the two nodes most responsible for planning tactical missions (Command 2 [WP5, 52, 53] and Control 1 [WP 88, 82, 83]) had the most Whiteboard files residing on their SC^4 systems.

Table 13

Number of Whiteboard Files Residing on Each Staff Member's Workstation by Mission

Position	Mission 2	Mission 3	Mission 4	Mission 5
WP6	0	0	0	0
WP62	4	5	4	3
WP63	1	1	2	1
WP5	4	1	3	5
WP52	9	3	4	3
WP53	0	1	1	1
WP88	4	1	6	6
WP82	5	3	5	2
WP83	1	1	2	2
WP84	0	1	0	1
WP99	3	1	2	3
WP92	5	2	3	3
WP93	0	0	1	1
WP94	0	0	0	0
Iron 6	0	1	2	1
Killer 6	0	1	3	1
Lightning 6	0	1	5	1
Mad Dog 6	1	2	2	3

Figure 17 shows the average number of Whiteboard files each node as well as the company commanders had for each mission. The data suggest that after Mission 2, the staff became more adept at ensuring that the same Whiteboard OPORD files were available in the nodes and to the company commanders.

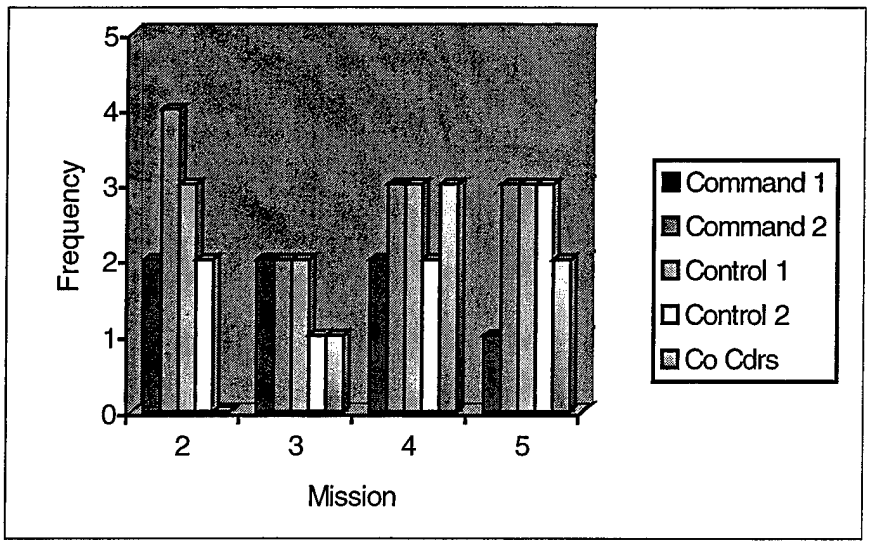


Figure 17. Average number of Whiteboard files for each node by mission.

Summary of Communication

The combination of three proposed automated measures (Whiteboard Use, Radio Communication Pattern, and Personnel Initiating Whiteboard Conferences) utilized during this project to provide insight into the skill dimension of Communication appeared somewhat useful, but all of these measures require further development. The Whiteboard Use measure, as designed, required strict adherence by the staff to file naming conventions. The Radio Communications Patterns, while providing some information in its present design, could benefit from voice recognition software to determine the extent to which staff members are communicating among themselves. The Information Flow measure may also require voice recognition software to determine the directions of radio and intercom communication to provide a clearer understanding of communications among the staff.

Coordination

Coordination is "the process by which team resources, activities, and responses are organized to ensure that tasks are integrated, synchronized, and completed within established temporal constraints," as indicated in Table 1. This may also be referred to as task organization, task interaction, or timing and activity pacing (Cannon-Bowers et al., 1995). The six candidate measures defined and presented to the SAG as ways to partially assess Coordination included Overlay Use, Whiteboard Commonality, Targeting, Fire Support Coordination, Fire Engagements, and OPFOR Destruction. The Project Team's premise was that these measures might suggest how well staff members coordinated their activities with each other.

Of these six candidate measures, all were developed. All conclusions regarding Coordination are based on these six measures. The Fire Support Coordination and Opposing Forces Destruction data are provided as examples of the output received from the Coordination measures implemented during BCR IV. Data from the other measures are provided in Appendix B.

Fire Support Coordination

The purpose of the Fire Support Coordination measure was to examine whether the staff had enough situational awareness to inflict damage on the OPFOR before they were within direct fire range. Table 14 shows the number of OPFOR kills with both direct and indirect fires. As seen in the table, the majority of the squadron kills were through direct fire for all four missions. Kills attributed to indirect fires were less than 50% for any given mission, and averaged around 30% per mission, which is considered good by current Army standards. However, these rates may be below average for what would be expected of a future digital force with more precise munitions.

Table 14

Frequency and Percentage of Opposing Forces (OPFOR) Kills Due to Direct and Indirect Fires by Mission

	Mission 2	Mission 3	Mission 4	Mission 5
Total OPFOR Losses	350	127	184	229
Total Battle Force Kills	125	7	45	75
Total Squadron Kills	225	120	139	154
Total Direct Fire Kills	127	92	111	88
Total Indirect Fire Kills	98	28	28	66
Total HIMARS (DS) Kills	57	28	22	65
Total 120Mor Kills	41	0	6	1
Total % Blue Indirect Kills	43%	23%	20%	43%
% HIMARS Kills	25%	23%	16%	42%
% 120Mor Kills	18%	0%	4%	1%

Note. HIMARS (DS) = high mobility artillery rocket system (direct support); 120Mor = 120mm Mortar Weapon System.

Figure 18 shows the number of direct and indirect fire kills for each mission. Missions 2 and 5 are quite close in the number of OPFOR killed by direct and indirect fires. However, Missions 3 and 4 have a much higher number of direct fire kills to indirect.

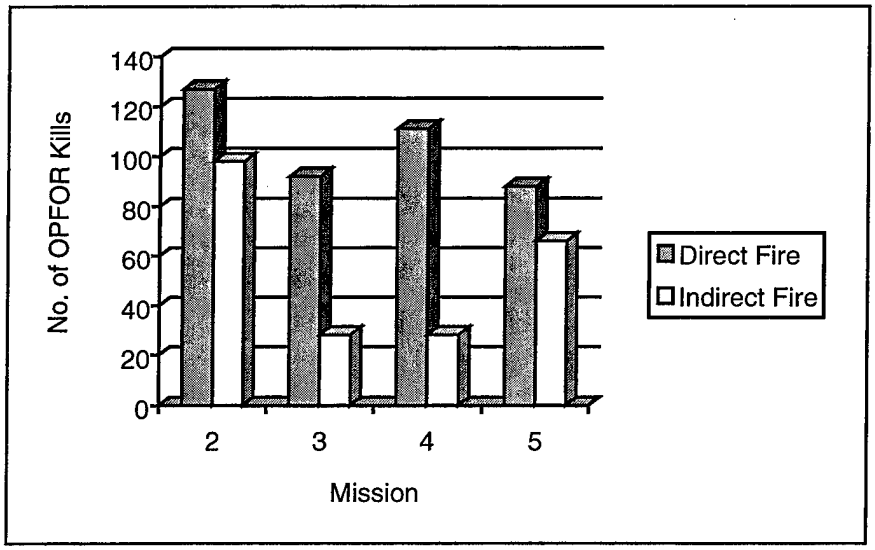


Figure 18. Frequency of opposing forces (OPFOR) kills due to direct and indirect fire for each mission.

The data appear skewed in this measure by the extended range of the direct fire weapon system that the squadron was equipped with. Some weapons systems, such as the squadron's FV, had both direct fire and beyond line of sight engagement capabilities. Because of this dual capability, the range differential between direct fire and indirect fire is probably decreased.

Opposing Forces Destruction

The OPFOR Destruction measure was designed to examine the battle tempo during the mission and an SC^4 equipped unit's ability to synchronize combat and combat support assets to efficiently destroy the OPFOR. Table 15 shows the cumulative rate of OPFOR losses for each mission. Losses are reported in 20 minute intervals, starting with 20 minutes after STARTEX.

Table 15

Cumulative Rate of Opposing Forces (OPFOR) Losses Every 20 Minutes for Each Mission

Time	OPFOR Losses			
	Mission 2	Mission 3	Mission 4	Mission 5
0:20	0	0	0	0
0:40	0	10	0	0
1:00	0	24	0	0
1:20	2	27	0	0
1:40	10	35	0	0
2:00	27	49	0	0
2:20	45	69	0	0
2:40	52	83	0	0
3:00	81	107	3	3
3:20	100	125	10	10
3:40	110	127	17	20
4:00	110		17	20
4:20	110		17	20
4:40	130		17	20
5:00	167		29	31
5:20	251		51	53
5:40	289		62	64
6:00	298		76	78
6:20	312		116	118
6:40	328		144	146
7:00	341		170	172
7:20	342		186	188
Total OPFOR	419	334	411	478

Figure 19 displays this same information in graph format. From the figure it becomes apparent that the rate of destruction in Missions 2 and 3 was quite similar, as was the rate of destruction in Missions 4 and 5. The rate data also reflect the distance between the two sides at the start of the mission. The closer the forces were at the start, the sooner engagement began, as seen in Missions 2 and 3. Once engagement began, the rate of destruction appears to be similar for each mission. This may mean that the squadron was uniformly efficient in attacking the OPFOR during each mission.

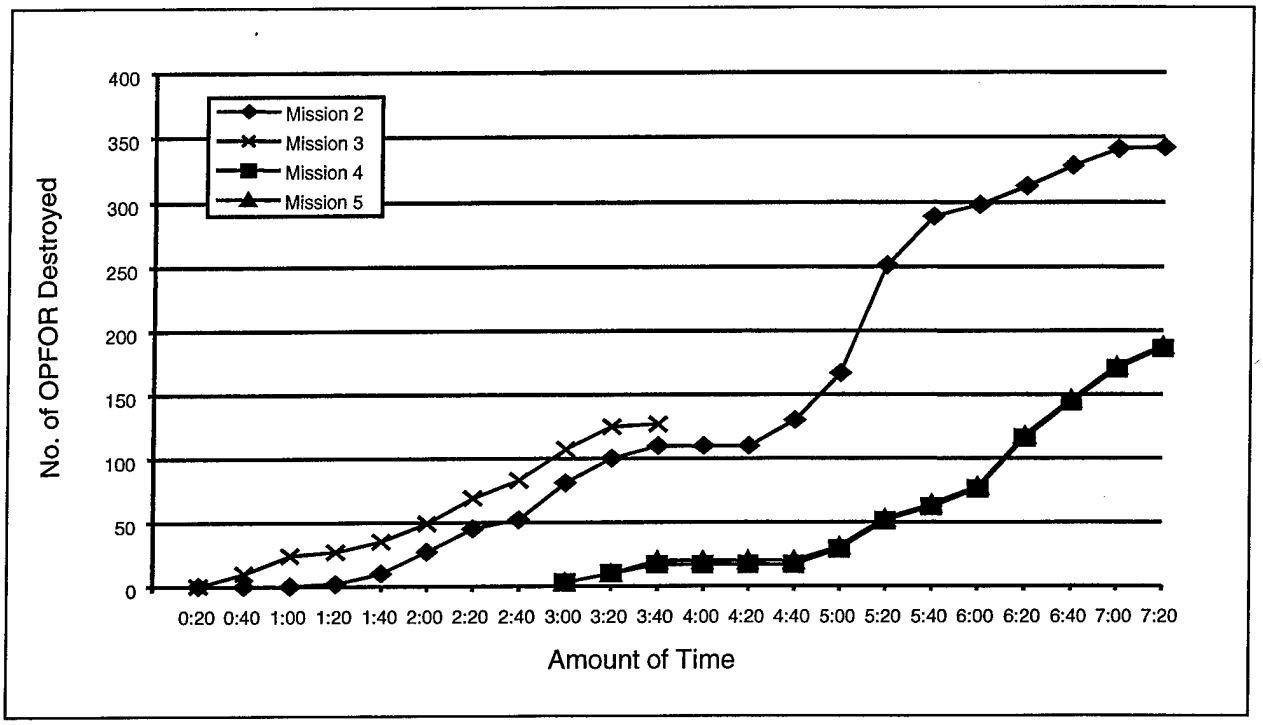


Figure 19. Cumulative rate of opposing forces (OPFOR) destruction for each mission over 20 minute time periods.

Summary of Coordination

The combination of automated measures Overlay Use, Whiteboard Commonality, Targeting, Fire Support Coordination, Fire Engagements, and OPFOR Destruction implemented during this project to provide insight into the skill dimension of Coordination appeared useful. The data obtained can be directly related to the synchronization, timely passing of relevant information, and team performance facilitation behaviors associated with this skill dimension. The measures developed by the Project Team for Coordination require little further development other than refinement of the output format to be useful for portraying Coordination.

Decision-Making

Decision-Making is "the ability to gather and integrate information, use sound judgment, identify alternatives, select the best solution, and evaluate the consequences," as indicated in Table 1. This may also be referred to as problem assessment, problem solving, planning, metacognitive behavior, or implementation (Cannon-Bowers et al., 1995). The four candidate measures defined and presented to the SAG as ways to partially assess Decision-Making included UAV Effectiveness, Length of Battalion Decision-Making Cycle (OPORD), Length of Battalion Decision-Making Cycle (Platoon Movement), and Information Retrieval by the Commander. The Project Team's premise was that UAV flight paths might serve as an indication of the underlying decisions on where to fly the UAVs. Additionally, shorter time intervals from receipt of higher headquarters OPORD until the squadron issues its own OPORD or until platoons begin to shift their positions might result in more favorable outcomes. The last measure sought to determine if the commander was getting all of the information required to

make a timely, correct decision. Of these four candidate measures, only UAV Effectiveness and Information Retrieval by the Commander were developed. All conclusions regarding Decision-Making are based on these two measures. The UAV Effectiveness data are provided as an example of the output received from the Decision-Making measures implemented during BCR IV. Data from the other measures are provided in Appendix B.

Unmanned Aerial Vehicle Effectiveness

This measure was designed to assess how the UAVs were being used, as a way to examine the underlying decisions that led to the UAV flight paths. The UAVs were directly controlled by the staff. If the UAVs were being used effectively, they might have been sent to areas of interest not already being covered by other sensors. This would provide sufficient early warning to allow the squadron to effectively and efficiently engage the OPFOR. Two output formats were designed for this measure: a summary table of how many and what percentage of OPFOR vehicles were detected by the UAVs (see Table 16), and a picture of the UAV routes for each mission and the location of OPFOR at first detection (see Figure 20). Table 16 reveals that overall, 47% of all OPFOR vehicles that the squadron's higher headquarters estimated it would oppose were first detected by the UAVs.

Table 16

Summary Table of Unmanned Aerial Vehicle (UAV) Effectiveness

Type Vehicle	Intelligence Estimate	Vehicles Detected by Squadron	% of Estimate Detected	Vehicles Detected by UAV	% Detected by UAV
Tanks (T90)	314	312	99%	147	47%
IFV (BMP2)	137	104	76%	66	63%
IFV (BMP3)	28	36	129%	7	19%
APC (BTR80)	270	246	91%	82	33%
APC (BRDM2K)	18	15	83%	2	13%
FA (ACRV)	106	13	12%	9	69%
FA (2S19)	349	81	23%	31	38%
FA (2S23)	58	57	98%	38	67%
FA (BM22)	264	78	30%	52	67%
ADA (2S6)	38	26	68%	16	62%
ADA (SA_13)	32	16	50%	9	56%
ADA (SA_15)	28	8	29%	4	50%
Totals	1642	992	60%	463	47%

Note. ADA = air defense artillery; APC = armored personnel carrier; FA = field artillery; IFV = infantry fighting vehicle.

As an example, Figure 20 shows the squadron's UAV flight paths for Mission 2. The enclosed area indicates the squadron area of operations. These UAV paths delineate only the area where the UAVs detected OPFOR not detected by any other sensors (e.g., ground sensors). Three of the four squadron UAVs that flew during Mission 2 were east of the area of operations, in the direction from which the OPFOR were attacking. The fourth UAV did not make it very

far before it either crashed or was shot down. The locations of the OPFOR vehicles at first detection by the squadron UAVs are also shown as "x" in the figure.

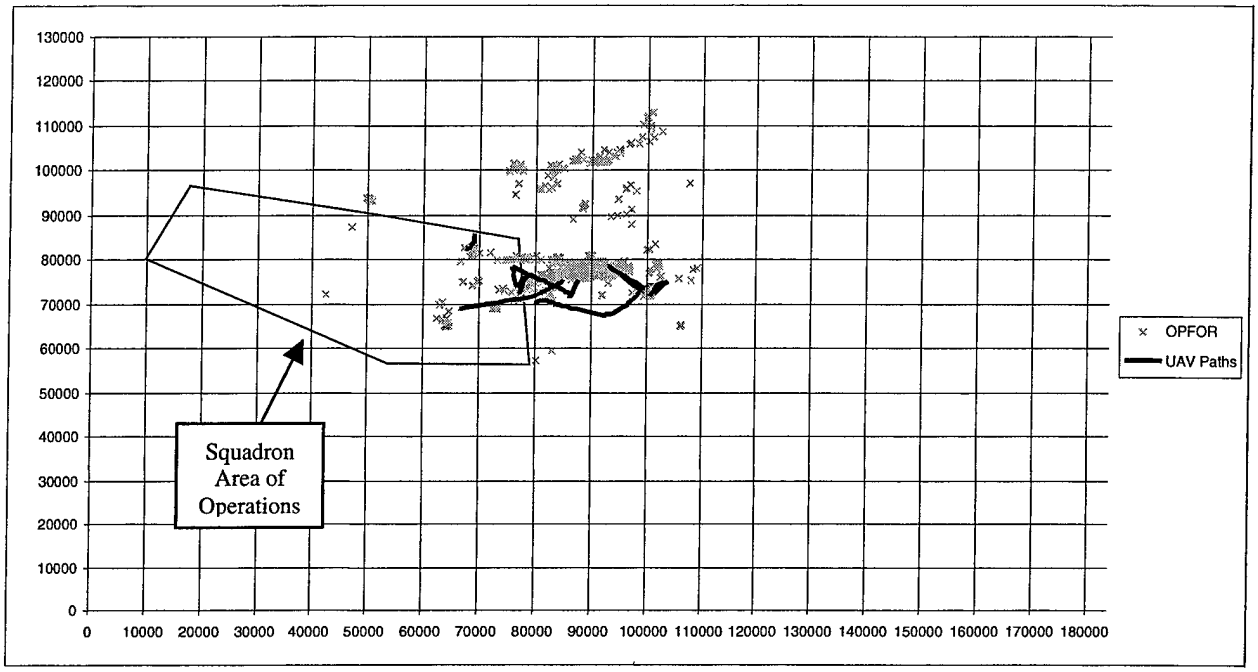


Figure 20. Unmanned aerial vehicle (UAV) flight routes and opposing forces (OPFOR) locations for Mission 2.

Summary of Decision-Making

The Project Team cannot make an overall assessment of utility based on the Decision-Making measures implemented. The UAV Effectiveness measure, especially the picture output, provided some useful information on the UAV flights, which shows where the OPFOR were first detected, but does not show the OPFOR that were missed which might provide more information on the decision-making process. Development on the two candidate measures not developed should be undertaken to determine if these types of measures can partially assess Decision-Making skills. The measures are designed to directly measure the length of time the commander and staff take to make a decision and disseminate it to subordinate commanders.

Summary

Overall, the project objective of developing and refining team and staff assessment strategies with a focus on automated measures was partially met. The Project Team refined, developed, and implemented a limited set of measures that have the potential, with further development, to become fully automated. The results obtained by these automated measures were not corroborated by any other data and, as mentioned earlier, should not be used to draw inferences about the squadron's performance during BCR IV. Many of the measures and formats developed, however, were used by the MMBL in support of the BCR IV evaluation

report (MMBL, in preparation). Nevertheless, much work remains to develop a more comprehensive, fully automated set of measures for assessing staff performance.

Battle Command Reengineering IV Implementation Lessons Learned

This project used the lessons learned mentioned in the Throne et al. (1999) report as a starting point for the development of the automated measures. That report identified the need for specifying the output format design at the same time that automated measures are being designed, before the measures themselves are developed. Although output formats were specified during this current project, other issues arose that needed to be addressed. Several lessons learned are presented here that may be of use to other researchers working in environments similar to the current MMBL setting. Table 17 contains a summary of these lessons learned.

Table 17

Summary of Implementation Lessons Learned

Lesson Learned	Summary
Automated Staff Performance Feedback	Intrinsic, or near real-time feedback to the staff on their collective performance based on the data from the automated measures is needed.
Measurement Framework	More research is needed to establish a viable measurement framework for assessing digital staff performance.
Balanced Measures	Automated measures are currently insufficient indicators of staff performance. Observers and training audience feedback is still required.
Measures Development	An iterative process for designing, developing, testing, and refining automated measures is required to get useful results.
Data Reduction	A plan for data reduction is essential due to the high volume of data collected in a digital environment.
Pictorial Output Formats	Development of pictorial representations is a worthwhile goal.
Automated Flags	A reliable list of automated flags is essential to data collection and analysis.

Automated Staff Performance Feedback

Additional research and development is required to provide near real-time feedback to a digital staff. One way this could be accomplished is by using data currently produced by constructive and virtual unit training simulations as the starting point for selected automated performance measures. Those data could then be posted to an Intranet web page so that unit personnel could review the results. An initial start to such a system was implemented during BCR IV by the MMBL programmer analysts.

Measurement Framework

The Project Team, early in the design process, adopted the Cannon-Bowers et al. (1995) model of team process skill dimensions as a starting point for the automated measures development since the experimental battalion staff organization in BCR IV did not have established standards on which to base measures. As described earlier, the experimental unit commander had complete flexibility in terms of how his staff would be organized, the role each staff member would perform, where they were located in the four C^2Vs dedicated to the staff, and what staff products he required to command and control his unit.

By adopting the Cannon-Bowers et al. (1995) model, a set of behaviors was identified as a measurement basis to assess staff performance. This approach was only partially successful. Six of the Cannon-Bowers et al. team process skill dimensions were used (see Table 6). The measures associated with the Coordination, Communication, and Shared Situational Awareness team process skill dimensions returned results that appear useful in assessing the experimental unit's staff performance in these areas. The measures appear useful because they may provide a direct and, objective (although partial) measure of performance in these team process skill dimensions. The Project Team is not as confident with the measures associated with Adaptability, Performance Monitoring and Feedback, and Decision-Making. While these measures provided useful results, they are less direct measures of performance, when compared to the Coordination, Communication, and Shared Situational Awareness measures. As a result, the Project Team was not able to assess the suitability of the measures in relation to these dimensions.

In general, the Cannon-Bowers et al. team process skill dimensions provided a generic and open-ended framework upon which to design prototype automated measures of team performance. This framework was reinforced by TRADOC's model of team skill requirements as well (DA, 1999). As the organization and functions of future brigade and battalion digital staffs become more fixed, specific staff tasks will be identified and baseline performance standards established which should result in more specific and direct measures of performance than the approach followed in this project.

Balanced Measures

Automated measures, regardless of how well they are designed, will not provide a complete assessment of staff performance. There continues to be a requirement for trained, experienced observers, and for feedback from the staff itself on their assessment of their own performance. While the complete set of automated measures designed during this project was not implemented during BCR IV, it is doubtful that a comprehensive assessment of staff performance could be made based on these automated measures without observer and unit inputs to balance the objective measures data. The increased scope and precision of data available from automated measures, however, should reduce observer and evaluation workload and subjectivity.

Measures Development

The majority of the prototype automated measures developed appeared useful in providing some feedback on the performance of the staff. This usefulness was reflected, in part, by the

MMBL's expressed support of the measures during a final project briefing and their request for the measures for their BCR IV evaluation report (MMBL, in preparation). Other candidate measures were not developed, due largely to BCR-related time and resource constraints. Future automated measures development should include the candidate measures not implemented during BCR IV to determine if they are useful for assessing staff performance as well as the team process skill dimensions they were designed to assess.

Nevertheless, all of the prototypes developed could be improved to provide more useful information. For example, the graphs developed for the Terrain Analysis measure would have been more useful if they provided the average frequency of use for each system tool instead of the average duration of each use (see Figures 7-10). Another potential graph would have been the percentage of time the tools were used in relation to the total mission time. A third alternative would have been to show the frequency and duration of use for each system tool in relation to critical events during the mission. These types of graphs may have provided more informative feedback than the current graphs developed for Terrain Analysis. The same applies to the PIR measure. Notably, technologies are eminently capable of transforming databases into a myriad of output formats.

Some of the prototype automated measures developed did not work as designed due to characteristics of the BCR environment. For example, the Collateral Damage measure (Table B-6, provided in Appendix B), which was designed to partially assess Shared Situational Awareness, was not very helpful due to the organization of the squadron. The squadron did not directly control indirect fire assets. All indirect fire weapons systems were controlled by company commanders. As a result, the responsibility for any collateral damage that occurred during a mission could not be attributed to an action of the staff which negated the value of the measure in terms of assessing the staff's Shared Situational Awareness.

In addition, some of the measures using Whiteboard files (see Table 13 and Figure 17) did not provide the expected information until an extensive manual review process was completed. The measures were developed based on the prediction that the squadron would establish an SOP for file naming conventions on critical information like OPORDs. In practice, the squadron did establish such an SOP, but did not consistently apply it. Consequently, the Project Team had to review each file manually to determine which files contained critical information that should have been distributed among the staff.

Finally, the Fire Support Coordination measure (see Table 14) did not provide a clear look at the contribution of direct fire and indirect fire support weapons systems to the outcome of the mission. Without this information, an assessment of the staff's influence on fire coordination could not be made. This shortcoming was created by addition of a hybrid gun system to the squadron weapons mix just prior to the start of the experiment. The hybrid gun, which was under the direct control of the company commanders, had both direct fire and indirect fire capabilities. The maximum range of the hybrid gun in the indirect fire mode was 12 times greater (25,000 versus 2,000 meters) than in the direct fire mode. This range capability exceeded all ranges for the other non-line of sight weapon systems in the squadron, including mortars and missiles. To account for this capability, each individual hybrid gun engagement would have to be examined to determine if in fact it was a direct fire engagement.

In a BCR-type setting, an iterative process for designing, developing, testing, and refining automated measures is required to get useful results. Since the programmers were in a different location from the SMEs and researchers, it was especially important to follow up on all measure outputs to make certain the data format met everyone's expectations. It is important that everyone involved in the data collection process have an understanding and common agreement of what the end product should be.

Data Reduction

Data reduction was a time consuming, iterative process for the BCR IV data analysis. During the design and development process, the Project Team had projected that various summary tables would provide the desired results with little additional processing. In analyzing the initial results, the Project Team determined that several measures (such as Radio Communications Pattern, OPFOR Destruction, and UAV Effectiveness) needed to have the supporting data tables refiltered or the basic data table used in lieu of a filtered table to get the desired data. A plan to reduce a significant amount of data is required if the initial measure design does not produce the desired result.

Pictorial Output Formats

The project's emphasis on development of pictorial formats for staff measures was only partially met. For example, showing the amount of battlefield each staff member displayed in a picture format appears much easier to understand than presenting the same information in a table (compare Table 12 to Figure 13). Only the squadron commander's and a company commander's views are presented in the figures presented in the report, but other staff members' views could be compared as well, based on the available measure data. The potential usefulness of pictorial representations as a form of staff performance feedback is a preliminary finding, since only 2 of the 13 candidate pictorial formats were actually developed.

Although development of the pictorial representations for the measures is a challenge, it appears a worthwhile goal. Even the limited set of prototype picture formats developed for this project appear promising. Pictorial formats not only effectively summarize a great deal of performance data, they also relate that performance directly to the temporal and spatial setting in which it occurred. Notably, the SAG and representatives from the MMBL reviewed the pictorial representations of the measures developed and expressed strong support for their potential for staff feedback and assessment.

Automated Flags

The potential of automated flags is evident. However, the full potential of implementing them to narrow data collection and analysis requirements for digital C^4I systems was not realized in this project. Of the 15 automated flags proposed during the development phase (see Table 7), only four were implemented after the initial examination the BCR IV trial data. The four utilized during the BCR IV data analysis were: STARTEX, first indirect fire engagement, first direct fire engagement, and last engagement. In most missions, the last engagement was also the end of exercise. The other nine automated flags were not implemented for a number of reasons. For example, the conditions for three automated flags (friendly losses exceeded 30%, OPFOR losses exceeded 70%, and command and control node loss) were not met during the BCR IV missions.

The importance of automated flags, however, is reinforced by the observer manual flag logs. In examining the manual logs maintained by the five observers, there was little commonality among them. If observers in close proximity, in the controlled environment of the BCR IV experiment could not maintain a manual log, it is most likely that manual logs would not be reliable when maintained under field conditions with widely dispersed command and control nodes. The need for this type of data collection and analysis bounding system was underscored by the Map Area measure designed and developed during this project. To obtain the results presented in this report (see Table 12 and Figure 13), approximately 18,000 records related to changes in map scale were examined and reduced to the 48 records needed to create the Map Area tables and figures for each mission.

It is very important to have a reliable list of automated flags. The flags should target events of interest to the researchers, as well as the training audience. Flags help researchers identify precisely what portions of the data they are interested in analyzing. In a BCR environment, the amount of data recorded during the missions is too extensive to analyze in its entirety. Without a way to filter the data, there will be too much information to analyze and present in a coherent fashion.

Summary

The implementation lessons learned during this project may assist other researchers developing automated measures of staff performance in the near future. Some of the insights that the Project Team gained while developing prototype automated measures of staff performance may have implications for future C^4I system designers and digital staff training developers. These are covered in the next section.

Future Efforts

Clearly, more research and development are needed before the Army can realize the potential of automated measures for training and performance evaluation. The prototype automated measures of staff processes described in this report are only a modest indication of the true potential of automated measures for more effective and efficient training and evaluation methods. Some of the future research and development that needs to be done to fully exploit the potential of automated measures is provided below. Table 18 contains a summary of these lessons learned for future efforts.

Table 18

Summary of Lessons Learned for Future Efforts

Lesson Learned	Summary
Measurement Framework	Measures need to be directly related to training and performance requirements.
Balanced Measures	Automated measures will become an increasingly important complement to traditional manual measures.
C^4I Instrumentation	Instrumented digital systems can, and should, provide a log of all soldier-computer interactions.
Digital Integration	The integrative nature of digital technologies, including C^4I and training systems, should be exploited to precisely correlate staff performance with battlefield conditions.
Multidisciplinary Teams	Multidisciplinary teams are essential to the development of useful automated measures.
Iterative Processes	The development of automated measures is an iterative process, building on knowledge gained with measure implementation and improvements to digital C^4I systems.
Data Flags	Logs need to be kept of key events and times during the exercise to collect and provide useful information to staff members.
Automated System Flags	Automated flags are essential to extract the desired data for real-time or near real-time feedback to staff members.
User-Defined Flags	Members of the training audience and O/Cs should be allowed to input flags into the automated measures system for events they consider important and want to discuss in an after action review.
Data Reduction	A plan for data reduction is essential due to the high volume of data collected in a digital environment.
Output Formats	Measures design should address output format requirements, and exploit the ability of digital technologies to provide more meaningful feedback, including pictorial formats.
Performance Feedback	Automated measures that provide both intrinsic and extrinsic feedback are needed.
Training Audience	The training audience is a key determinant of useful measures and meaningful formats, and should be included in all measurement development efforts.

Note. C^4I = command, control, communications, and computers; O/C = observer/controller.

Measurement Framework

When designing and developing automated measures for digital staff training and evaluation, it is important to have a measurement framework that relates measures to training and performance requirements. If the current research on the Army's future battalion and brigade-level battle staff results in a smaller, geographically dispersed staff operating advanced C^4I systems, a conventional approach to training and performance assessment that relies primarily on observers, which is the case today, may not provide meaningful feedback. New approaches will have to be developed and fielded concurrently with the C^4I systems provided to the staffs.

Balanced Measures

Traditional methods for assessing C^2 performance, including staff performance, are heavily "burdened" measures (i.e., resource intensive) and highly subjective. The measurement burden is further compounded by the inherent complexity of C^2 performance including the requirement to track and integrate communication exchanges across numerous and dispersed participants (Crumley, 1989; Olmstead 1992). As a result, traditional staff performance measurement methods (e.g., surveys and observations) often fail to provide a detailed and objective database for adequately assessing staff performance. Moreover, the workload and time lags required for manual integration and interpretation of staff performance data obtained by traditional methods often prevent timely and meaningful feedback to the training audience.

The more staffs rely on computers to do their work, the more their computer interactions will become meaningful and measurable aspects of work process and products. Automated measures are not only objective, they are direct measures *of* performance. In contrast, many traditional measurement methods (e.g., survey, interview, observation) are measures *about* performance. The ability of digital technologies to track soldier and computer interaction should increase the scope and precision of performance assessment and feedback. In addition, unobtrusive and automatic data collection may reduce measurement error as well as observer workload and resource requirements. Perhaps, most importantly, automated measures correlate performance to the operational setting in which it occurs.

However, no single measurement method or measure is sufficient. While automated measures, might provide the preponderance of objective data about staff performance, complementary manual measurement approaches are still needed from trained observers and from the staff itself. Tools need to be developed that can allow these inputs to be fused into the training and performance assessment provided to the unit. By employing a balanced set of measurement methods, trainers and researchers can achieve a fuller understanding of staff performance, and staffs can get the feedback they need to strengthen their performance

Command, Control, Communications, Computers, and Intelligence Instrumentation

Any set of automated measures of staff performance will depend on instrumenting and linking individual soldier C^4I workstations, combat systems, and strategic and tactical C^4I systems that are feeding information to the collective staff or to individual staff members. Researchers have noted, in a simulation context, that there is a divergence of information flow

that creates a challenge in providing digital training feedback. There are four types of Information Flow: C^4I system digital data, simulation data, O/C data, and test data. The C^4I system and simulation data are generally on separate information networks from the test data network, while the O/C data is often manually recorded (Gerlock & Meliza, 1999). The situation just described could be the one faced by future trainers of staffs equipped with complex, integrated digital C^4I systems if workstations, combat systems, and information networks are not instrumented to collect data about their use and status, and linked with one another to share these data.

The MMBL's DCA, which is unique within the Army, solves this data integration problem and takes it a step further by providing a means to link performance data to the tactical setting in which it occurred. Recent improvements to the DCA allow soldiers to use a browser to view various reports and data tables that have been posted to an Intranet web page during training and exercises. This approach provides timely feedback to the soldiers on the unit's performance and may be applicable to future efforts to provide performance feedback. Overall, the DCA provides an approach to data collection and analysis on which future efforts to automate performance measurement of staffs equipped with digital C^4I systems can be based. Adding voice recognition and graphing format capabilities would increase the utility of the DCA for automated performance measurement and feedback.

Digital Integration

Automated measures of staff performance can and should provide performance data related to the tactical context in which performance occurred. The technology that records staff performance data is also recording unit actions and status. Individual soldier-SC^4 system interactions are also being recorded. Given the ability of digital technology to correlate temporal and spatial dimensions across systems, these different recording systems can and should be integrated to provide a common picture of the battlefield upon which to analyze staff performance.

Multidisciplinary Teams

The use of multidisciplinary teams is essential to the development of useful automated measures of staff performance. At a minimum, researchers, SMEs, and digital C^4I system programmers are needed to identify potential measures of staff performance and then design, develop, and implement them. Unfortunately, system programmers are usually included during the measures development stage, which may be too late. Including programmers during the measures design process will allow researchers and SMEs to determine what types of measures are feasible given the digital C^4I system limitations. The synergy of this team may also result in new specifications and formats.

Iterative Processes

The design and development of measures by multidisciplinary teams is an iterative process. Computer hardware configurations and software programs may initially define the boundaries of staff performance that can be measured automatically. As the operational definitions and output formats of the measures are conceptualized and the development processes are initiated, changes

to computer hardware and software programs may be required which could lead to further measures analysis. As long as digital C^4I capabilities keep improving and becoming more integral to staff performance, automated measures development needs to keep pace or risk becoming meaningless for performance feedback.

Data Flags

Tactical training exercises are usually quite lengthy and involved. In order to collect and provide useful information to the staff or unit regarding training, an event-based log needs to be kept of what happened during the exercise and when. Otherwise, key events may be missed or not covered during feedback sessions. For staffs equipped with digital C^4I systems, there need to be two methods to determine when relevant data need to be collected: automated system flags and user-defined flags.

Automated System Flags

Not only are automated flags important for C^4I data analysis, they are essential. Without automated flags to identify critical points during an exercise, the amount of data will be too much to present meaningful results. If the system can automatically keep track of when key events such as first direct fire occur, then the output data for any given amount of time before or after that event took place can be examined. Future automated measures of staff performance will need these flags to extract the desired data if real-time or near real-time feedback to staff members is to be achieved. Staff training developers should continue to examine the concept and refine automated flags for measures design and development.

User-Defined Flags

A process is required that allows members of the training audience or an O/C to input an event flag into the DCA system so that automated measurement routines would recognize the event and include it in the results that are automatically produced at the end of training. This capability would enhance staff training by allowing staff members to mark events that they consider noteworthy or unusual, obtain data related to the events in the context in which they are occurring, and to have the data available at the end of training for analysis and discussion among the rest of the staff. This capability would also provide the O/Cs with flexibility in recording data about an unexpected event.

Data Reduction

As a staff's digital capabilities advance and staff trainers using constructive and virtual simulation become more widespread, training exercises will become more detailed and complex. This will lead to an exponential increase in data with potentially useful information that can assist in assessing staff performance. Therefore, a plan for data reduction is essential. As previously discussed, using a system of flags, automated and user-defined, can focus attention on essential data, which should help in the data reduction. In terms of future automated performance feedback measures design, the data reduction and filtering issues require extended research and development efforts. Fully automated measurement will require extensive filtering

routines to ensure that programmers are able to exclude all data not directly related to the measure.

Output Formats

Output formats of automated staff performance measures need to be identified during the measures design process. Researchers, SMEs, and system programmers need to work together to identify the most useful formats for presenting the data obtained for each measure. Even important findings will lose their usefulness if they are not presented in a meaningful format.

Performance Feedback

Based on their own observations of staff performance, Kirlik, Fisk, Walker, and Rothrock (1998) concluded that feedback provided to trainees could be improved in four areas. First, the timeliness of feedback could be improved so that trainees do not have to wait until the AAR to receive feedback. Even when trainees do receive on-line or intrinsic feedback, it usually occurs after a critical error has taken place, by which point it may be too late to modify the behaviors that led to the occurrence. Second, feedback needs to be standardized so that O/Cs use a consistent approach instead of developing their own priorities for what is most important. Third, although it may be useful to measure and report trainees' frequency of errors, it is more useful to link the errors to their consequences and to lessons learned. Fourth, some trainees viewed verbal feedback during training as an interruption to the tasks they were trying to accomplish. Feedback that strikes a balance between enhancing training effectiveness and being unobtrusive to the trainee is needed.

Digital C^4I systems have the potential, by means of automated measures, to provide improved intrinsic and extrinsic feedback to the staff. Intrinsic feedback can be as overt as telling the individual staff member that he is not entering the correct information that a particular C^4I system tool needs to perform an action, or as subtle as the display of the tactical movement of a subordinate unit in response to a fragmentary order issued by the staff. Extrinsic feedback can be provided directly from the output of automated measures during the AAR process.

Training Audience

Training audience representatives should have input into the development process for automated measures of performance feedback. Their insights can help determine what should be measured and, perhaps more importantly, provide recommendations on how to display the data produced by a measure so that it can be readily understood by soldiers. After the automated performance measures have been developed and implemented, the training audience should be able to select, during their training, a subset of measures that focus on the areas the commander and staff feel that they need feedback in order to improve their performance. A fixed set of automated measures may be too encompassing. The fixed set could overwhelm the staff with data that tell them what they already knew, in which case it would be disregarded. The fixed set might also have been too narrowly defined, in which case a process needs to be in place that would allow the training audience to define their own measures, specify output formats, and obtain the results without having to go through a design and development process.

Summary

As the Army reshapes itself into an information-age force, with Future Combat Systems that are lighter, more lethal, and fully integrated into the C^4I environment, trainers must begin to anticipate the changing collective task demands and training requirements that will confront future (e.g., 2012 and beyond) soldiers and leaders. These soldiers and leaders are likely to be multi-functional, responsible for a myriad of tasks as yet undefined. This future training environment may rely heavily on technology to provide distributed and embedded training approaches to meet collective and unit-based training requirements, including command and staff training.[5]

The operating environment of future battle staffs will also evolve to take advantage of the inherent capabilities of advanced digital C^4I systems. Battle staffs may be smaller, operating in specialized C^2Vs, and geographically dispersed across a non-linear battlefield. The tasks traditionally associated with staffs may be changed dramatically as well. Expert computer systems employing artificial intelligence software may handle most routine analysis, estimation, and information distribution tasks associated with battle staffs. Considerable research will be needed to formulate performance standards for future staffs. An approach like the team process skills framework adopted during this project may complement more traditional measures of staff products such as orders and estimates.

Closely linked to the training requirement is the need for measurement, both to allow for feedback and performance improvement, and also to support the design and development of the training programs. Traditional training measurements that rely on direct observation and objective measurement of performance during training are difficult, particularly for C^2 performance. Subjective methods used for assessing C^2 performance are labor-intensive approaches, requiring observers with high subject matter expertise. Even with automated data collection aids such as electronic clipboards or computer-assisted observation tools, these methods are inefficient. They are also subject to unreliability if there is a lack of standardization among observers. This common problem is further exacerbated in the information-intensive environment of digital C^4I systems. Observers, like users and participants, can be quickly overwhelmed with the amount of information relevant to C^2 performance.

Automated measures of performance may well solve some of the challenges associated with employing traditional training measurement methods in a digital environment, but they will not, by themselves, be able to describe all facets of performance. These automated measures, however, may collect the preponderance of performance data required for meaningful feedback. Future research is needed to establish a balance among automated measures, observations, and training audience feedback through an AAR process.

As C^4I systems become more integral to the performance of individual and collective tasks, the human-computer interactions associated with these systems become more critical and collectible. The digital advances in Army units have increased the need for increased training in information management and situational awareness skills. These digital advances also increase

[5] Two recent ARI-sponsored research projects have explored prototype training methods for future staff that could be a viable training model. For more information, see Throne et al. (1999) or Deatz et al. (in preparation).

the need for timely, objective feedback of training and performance. Digital C^4I systems have unrealized potential to capture and analyze data on soldier, small-group, and collective performance. Future force requirements and capabilities underscore the need for additional research and development on automated measures of performance that provide intrinsic and extrinsic feedback to both soldiers and battle staffs.

References

Atwood, N. K., Quinkert, K. A., Campbell, M. R., Lameier, K. F., Leibrecht, B. C., & Doherty, W. J. (1991). Combat vehicle command and control systems: Training implications based on company-level simulations (ARI Technical Report 943). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Baker, D. P., & Salas, E. (1992). Principles for measuring teamwork skills. Human Factors, 34, 469-475.

Baker, D. P., & Salas, E. (1997). Principles for measuring teamwork: A summary and look toward the future. In M. T. Brannick, E. Salas, & C. Prince (Eds.), Team performance assessment and measurement (pp. 331-355). Mahwah, NJ: Lawrence Erlbaum Associates.

Brannick, M. T., Prince, A., Prince, C., & Salas, E. (1995). The measurement of team process. Human Factors, 37(3), 641-651.

Brown, B., Nordyke, J. Gerlock, D., Begley, I. & Meliza, L. (1998). Training analysis and feedback aids (TAAF Aids) study for live training support (ARI Study Report 98-04). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Caldera, L., & Reimer, D. J. (1999). A statement on the posture of the United States Army fiscal year 2000 [On-line]. Available: http://www.army.mil/aps/00/aps00.htm.

Cannon-Bowers, J. A., & Salas, E. (1997). A framework for developing team performance measures in training. In M. T. Brannick, E. Salas, & C. Prince (Eds.), Team performance assessment and measurement (pp. 45-62). Mahwah, NJ: Lawrence Erlbaum Associates.

Cannon-Bowers, J. A., Tannenbaum, S. I., Salas, E., & Volpe, C. E. (1995). Defining competencies and establishing team training requirements. In R. Guzzo & E. Salas (Eds.), Team effectiveness and decision making in organizations (pp. 333-380). San Francisco, CA: Jossey-Bass Publishers.

Cooper, M., Shiflett, S., Korotkin, A. L., & Fleishman, E. A. (1984). Command and control teams: Techniques for assessing team performance (AFHRL-TP-84-3). Brooks Air Force Base, TX: Air Force Human Resources Lab.

Crumley, L. M. (1989). Review of research and methodologies relevant to army command and control performance measurement (ARI Technical Report 825). Alexandria, VA: U.S. Army Research Institute for the Social and Behavioral Sciences.

Deatz, R. C., Greene, K. A., Holden, W. T. Jr., Throne, M. H., & Lickteig, C. W. (in preparation). Refinement of prototype staff training methods for future forces (ARI Research Report). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Department of the Army. (1997a). Field Manual 101-5, Staff organization and operations. Washington, DC: Headquarters Department of the Army.

Department of the Army. (1997b). Field Manual 101-5-1, Operational terms and graphics. Washington, DC: Headquarters Department of the Army.

Department of the Army. (1999). TRADOC Regulation 350-70, Systems approach to training management, processes, and products. Fort Monroe, VA: Headquarters, U.S. Army Training and Doctrine Command.

Dwyer, D. J., Fowlkes, J. E., Oser, R. L., Salas, E., & Lane, N. E. (1997). Team performance measurement in distributed environments: The TARGETs methodology. In M. T. Brannick, E. Salas, & C. Prince (Eds.), Team performance assessment and measurement (pp. 137-153). Mahwah, NJ: Lawrence Erlbaum Associates.

Dwyer, D. J., Oser, R. L., Salas, E., & Fowlkes, J. E. (1999). Performance measurement in distributed environments: Initial results and implications for training. Military Psychology, 11(2), 189-215.

Dzindolet, M. T., Gleason, D. R., Hernandez, C., Hill, K., Jaquint, M., Jette, F., Larsen, J., Oxley, N. L., & Pierce, L. G. (1998). Defining training objectives and performance measures in a Janus battle simulation (ARL-CR-433). Aberdeen Proving Ground, MD: U.S. Army Research Lab.

Fowlkes, J. E., Lane, N. E., Salas, E., Franz, T., & Oser, R. (1994). Improving the measurement of team performance: The TARGETs methodology. Military Psychology, 6(1), 47-61.

Gerlock, D. L., & Meliza, L. L. (1999). Supporting exercise control and feedback in the digital domain for virtual simulations. Proceedings of the 41st Annual Conference of the International Military Testing Association (CD-ROM). Monterey, CA: International Military and Testing Association.

Hayes, R. E., Layton, R. L., & Ross, W. A. (1995). The Army Command and Control Evaluation System documentation (ARI Research Product 95-09). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Johnston, J. H., Smith-Jentsch, K. A., & Cannon-Bowers, J. A. (1997). Performance measurement tools for enhancing team decision-making training. In M. T. Brannick, E. Salas, & C. Prince (Eds.), Team performance assessment and measurement: Theory, methods, and applications (pp. 311-327). Mahwah, NJ: Lawrence Erlbaum Associates.

Kirlik, A., Fisk, A. D., Walker, N., & Rothrock, L. (1998). Feedback augmentation and part-task practice in training dynamic decision-making skills. In J. A. Cannon-Bowers & E. Salas (Eds.), Making decisions under stress: Implications for individual and team training (pp. 91-113). Washington, DC: American Psychological Association.

Leibrecht, B. C., Meade, G. A., Schmidt, J. H., Doherty, W. J., & Lickteig, C. W. (1994). Evaluation of the combat vehicle command control systems: Operational effectiveness of an armor battalion (ARI Technical Report 998). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Lickteig, C. W., & Collins, J. W. III. (1995). Combat vehicle command and control system evaluation: Vertical integration of an armor battalion (ARI Technical Report 1021). Alexandria, VA: U.S. Army Research Institute for the Social and Behavioral Sciences.

Lickteig, C. W., & Throne, M. H. (1999). Applying digital technologies to training: A focus on pictorial communication (ARI Technical Report 1097). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Mason, D. E. (1995). Identifying measures of effectiveness for Marine Corps C^4I systems. (Master's Thesis). Monterey, CA: Naval Postgraduate School.

McIntyre, R.M., & Salas, E. (1995). Measuring and managing for team performance: Emerging principles from complex environments. In R. Guzzo & E. Salas (Eds.), Team effectiveness and decision making in organizations (pp. 9-45). San Francisco, CA: Jossey-Bass Publishers.

Mohs, B., MacDiarmid, W. R., & Andrews, D. H. (1988). A review and annotated bibliography of training performance measurement and assessment literature (ARI Technical Report 815). Alexandria, VA: U.S. Army Research Institute for the Social and Behavioral Sciences.

Mounted Maneuver Battlespace Lab (1998). Battle lab experiment plan (BLEP) for battle command reengineering III (BCR III). (Available from Mounted Maneuver Battlespace Lab, Attn: ATZK-MW, Fort Knox, KY 40121)

Mounted Maneuver Battlespace Lab (1999). Battle lab experiment final report (BLEFR) for battle command reengineering, phase III. (Available from Mounted Maneuver Battlespace Lab, Attn: ATZK-MW, Fort Knox, KY 40121)

Mounted Maneuver Battlespace Lab (in preparation). Battle lab experiment final report (BLEFR) for battle command reengineering, phase IV. (Available from Mounted Maneuver Battlespace Lab, Attn: ATZK-MW, Fort Knox, KY 40121)

Olmstead, J. A. (1992). Battle staff integration (IDA Paper P-2560). Alexandria, VA: Institute for Defense Analyses.

Oser, R., McCallum, G. A., Salas, E., & Morgan, B. B. Jr. (1989). Toward a definition of teamwork: An analysis of critical team behaviors (Technical Report 89-004). Orlando, FL: Naval Training Systems Center.

Salas, E., Bowers, C. A., & Cannon-Bowers, J. A. (1995). Military team research: 10 years of progress. Special Issue: Team processes, training, and performance. Military Psychology, 7, 55-75.

Salas, E., Dickinson, T. L., Converse, S., Tannenbaum, S. I. (1992). Toward an understanding of team performance and training. In R. W. Swezey and E. Salas (Eds.), Teams: Their training and performance (pp. 3-29). Norwood, NJ: Ablex Publishing Company.

Smith-Jentsch, K. A., Johnston, J. H., & Payne, S. C. (1998). Measuring team-related expertise in complex environments. In J. A. Cannon-Bowers & E. Salas (Eds.), Making decisions under stress: Implications for individual and team training (pp. 61-87). Washington, DC: American Psychological Association.

Sterling, B. S., & Quinkert, K. A. (1998). Review of battle staff training research at brigade and battalion levels (ARI Research Report 1730). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Throne, M. H., Deatz, R. C., Holden, W. T. Jr., Campbell, C. H., Sterling, B. S., & Lickteig, C. W. (1999). Prototype staff training and evaluation methods for future forces (ARI Research Report 1745). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

U.S. Army Research Institute for the Behavioral and Social Sciences (1997). The U.S. Army Research Institute's Science and Technology (S&T) program for FY 98. Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

U.S. Army Research Institute for the Behavioral and Social Sciences (1999). Training and Measurement Support Package, Battle Command Reengineering III, Mounted Maneuver Battlespace Lab, Volumes 1-5. (Available from Mounted Maneuver Battlespace Lab, Attn: ATZK-MW, Fort Knox, KY 40121)

U.S. Army Research Institute for the Behavioral and Social Sciences (2000). Training and Measurement Support Package, Battle Command Reengineering IV, Mounted Maneuver Battlespace Lab, Volumes 1-6. (Available from Mounted Maneuver Battlespace Lab, Attn: ATZK-MW, Fort Knox, KY 40121)

Appendix A

List of Acronyms

120Mor..... 120mm Mortar Weapon System

AAR..... after action review
ABF..... attack by fire
ACCES..... Army Command and Control Evaluation System
ADA..... air defense artillery
APC..... armored personnel carrier
ARI..... U.S. Army Research Institute for the Behavioral and Social Sciences
ARPA..... Advanced Research Projects Agency
ARSI..... ARPA Reconfigurable Simulator Initiative

BCR..... Battle Command Reengineering
BCV..... battle command vehicle
BLEFR..... Battle Lab Experiment Final Report
BLUFOR..... blue forces
BOS..... battlefield operating system

C^2..... command and control
C^2V..... command and control vehicle
C^4I..... command, control, communications, computers, and intelligence
CCIR..... commander's critical information requirements
CEP..... Concept Experimentation Program
CVCC..... Combat Vehicle Command and Control

DA..... Department of the Army
DC^4I..... Prototype Methods for the Design and Evaluation of Training and Assessment of Digital Staffs and Crewmen
DC^4I-2..... Refinement of Methods for the Training and Assessment of Digital Staffs
DCA..... Data Collection and Analysis System
DI..... dismounted infantry
DIS..... distributed interactive simulation
DW..... displayed world

ENDEX..... end of exercise

FA..... field artillery
FASTRAIN..... Force XXI Training Methods and Strategies
FBC..... future battlefield conditions
FLOT..... forward line of troops
FOV..... field of view
FRAGO..... fragment order
FV..... fighting vehicle

HIMARS (DS) .... high mobility artillery rocket system (direct support)
HumRRO ..... Human Resources Research Organization

IFV ..... infantry fighting vehicle
IITRI..... Illinois Institute of Technology Research Institute
INTEL ..... Intelligence Report

KL ..... kill

LOS ..... line of sight

MCOO..... Modified Combined Obstacle Overlay
MUAV ..... micro-UAV
MMBL ..... Mounted Maneuver Battlespace Lab
ModSAF..... Modular Semi-Automated Forces
MWTB ..... Mounted Warfare Test Bed

NCO ..... non-commissioned officer
NLOS ..... non-line of sight

O/C..... observer/controller
OCOKA ..... observation, cover, obstacles, key terrain, and avenues of approach
OIC..... officer in charge
OPFOR..... opposing forces
Ops ..... operations

PDU..... protocol data units
PIR ..... priority information request
PVD..... plan view display

RBC..... reengineered battle command
RW ..... real world

SA ..... situational awareness
SAG..... Subject Matter Expert Advisory Group
SBF..... support by fire
SC^4 ..... surrogate command, control, communications, and computers
SITREP ..... situation report
SME ..... subject matter expert
SOP ..... standing operating procedure
SOV..... staff operations vehicle
SPOTREP ..... spot report
STARTEX..... start of exercise
STO ..... Science and Technology Objective
SUV..... sport utility vehicle

SW..... sensed world

TARGETs ..... targeted acceptable responses to generated events or tasks
TRADOC ..... U.S. Army Training and Doctrine Command
TTP ..... tactics, techniques, and procedures

UAV ..... unmanned aerial vehicle
UGV ..... unmanned ground vehicle
USITREP ..... unit situation report

VSITREP ..... vehicle situation report
VTC..... video teleconference
VW ..... visible world

Appendix B

Prototype Automated Measurement Package

This appendix contains the operational definition, rationale, and recommended output format for each of the candidate automated measures. For those measures that were implemented during Battle Command Reengineering (BCR) IV, sample outputs are provided unless they were used as samples in the main body of the report.

Candidate Automated Measures Categorized by Team Process Skill Dimensions

Adaptability

1. Terrain Analysis
   a. Operational definition: Record the amount of time each node position uses and/or initiates each of the following tools: the Stealth Control, Terrain Intervisibility Tool, Field of View (FOV) Tool, Snail Display, and Forward Line of Troops (FLOT) Display during the mission.
   b. Rationale: Provides insight into whether the various terrain aids and other tools available in the surrogate command, control, communications, and computers (SC^4) system assist the commander and planning staff in obtaining a better appreciation of the effects of terrain (observation, cover, obstacles, key terrain, and avenues of approach [OCOKA]) on the mission. Measures the behaviors related to having back-up plans and adjusting quickly to situational change.
   c. Recommended output formats:
      1) Table – Frequency and duration of use of Stealth Control, Terrain Tool, FOV Tool, Snail Display, and FLOT Display by mission by position
      2) Graph – Average duration of use of above tools by mission (bar graph)
   d. Data:
      1) Graphs and sample table provided in the main body of the report.
      2) Additional tables provided below.

Table B-1

Frequency and Duration of the Field of View Tool Use for Each Mission by Position[1]

Position	Mission 2		Mission 3		Mission 4		Mission 5	
	Frequency	Duration	Frequency	Duration	Frequency	Duration	Frequency	Duration
WP62	0	0:00	1	0:03	1	0:11	0	0:00
WP52	1	0:05	2	0:26	8	13:29	4	1:11
WP53	2	0:29	1	0:05	2	1:48	0	0:00
WP88	0	0:00	0	0:00	2	2:26	0	0:00
WP82	0	0:00	12	13:07	1	1:08	1	0:02
WP83	4	2:59	1	1:09	4	2:44	0	0:00

[1] Durations are given in minutes and seconds.

Table B-2

Frequency and Duration of the Snail Tool Use for Each Mission by Position[2]

Position	Mission 2		Mission 3		Mission 4		Mission 5	
	Frequency	Duration	Frequency	Duration	Frequency	Duration	Frequency	Duration
WP53	3	0:15	0	0:00	1	0:29	0	0:00
WP82	3	1:25	0	0:00	0	0:00	0	0:00
WP83	0	0:00	0	0:00	0	0:00	6	1:48
WP92	0	0:00	0	0:00	0	0:00	9	0:24

Table B-3

Frequency and Duration of the Stealth Tool Use for Each Mission by Position[3]

Position	Mission 2		Mission 3		Mission 4		Mission 5	
	Frequency	Duration	Frequency	Duration	Frequency	Duration	Frequency	Duration
WP6	2	4:00	2	2:35	0	0:00	0	0:00
WP63	1	0:47	4	9:27	9	9:51	1	44:38
WP5	8	2:30:00	1	2:29:24	0	0:00	14	1:37:21
WP53	0	0:00	0	0:00	0	0:00	10	2:19
WP83	0	0:00	1	0:19	0	0:00	0	0:00
WP99	9	2:54	0	0:00	3	0:44	2	0:18

2. Node Location
   a. Operational definition: Calculate the average distance of each node from each other and from the major combat units in the battalion at critical points during the mission (first indirect fire engagement with opposing forces (OPFOR); first direct fire engagement with OPFOR; first friendly casualty; friendly losses exceed 30%; and last engagement during mission).
   b. Rationale: The dispersion of the battalion's command and control nodes may indicate the ability of the battalion staff to handle different requirements simultaneously while keeping positioned to maintain communications with all subordinate elements and maintaining operational and physical security. Measures the behaviors related to having back-up plans, transitioning smoothly to back-up, and adjusting quickly to situational change.

[2] Durations are given in minutes and seconds.
[3] Durations are given in hours, minutes, and seconds.

c. Recommended output formats:
   1) Table – Average distance of each node from each other and from the major combat units in the battalion at critical points during the mission by node by mission.
   2) Picture – Snapshot of where nodes are located on battlefield at critical points during the mission.

d. Data:
   1) This measure was not implemented during BCR IV.

3. Loss of Node
   a. Operational definition: Calculate the average distance of each node from each other and from the major combat units in the battalion at the time of a loss of one or more nodes and at one hour after the loss of one or more nodes.
   b. Rationale: The relocation of the battalion's command and control nodes may indicate the ability of the battalion staff to adapt to loss of a command and/or control node. Measures the behaviors related to having back-up plans, transitioning smoothly to back-up, adjusting quickly to situational change, performing tasks outside job when asked, and changing the way task performed when asked.
   c. Recommended output formats:
      1) Table – Average distance of each node from each other and from the major combat units in the battalion at the time of a loss of one or more nodes and at one hour after the loss of one or more nodes by node by mission.
      2) Picture – Snapshot of where nodes are located on battlefield at the time of a loss of a node and at one hour after the loss of that node for each loss.
   d. Data:
      1) This measure was not implemented during BCR IV.

Performance Monitoring and Feedback

1. Situation Report (SITREP) Use
   a. Operational definition: Calculate the frequency, type (individual vehicle or unit) and duration of use of both types of SITREP tools during the mission by node position. By clicking on a vehicle picture on the plan view display (PVD), a staff member would get information on that individual vehicle situation report (VSITREP), such as fuel and ammunition status. By clicking on a unit icon on the PVD, a staff member would get information on all the individual vehicles in that unit situation report (USITREP). Additionally, record the frequency and duration of the "Auto On" feature and the frequency of the "detail" feature. The data should be collected for each staff position.
   b. Rationale: Data may indicate the frequency of the staff's monitoring of activities and unit status during a mission. Measures the behaviors related to responding to requests for information, proving constructive suggestions, and obtaining information about the outcomes of decisions and actions where appropriate.
   c. Recommended output formats:
      1) Table – Frequency, type, and duration of SITREP use by mission by position
      2) Graph – Frequency of SITREP use by mission by position (line graph)
      3) Graph – Average frequency of SITREP use by mission (bar graph)

d. Data:
   1) Sample bar graph provided below.
   2) Table and line graph not produced.

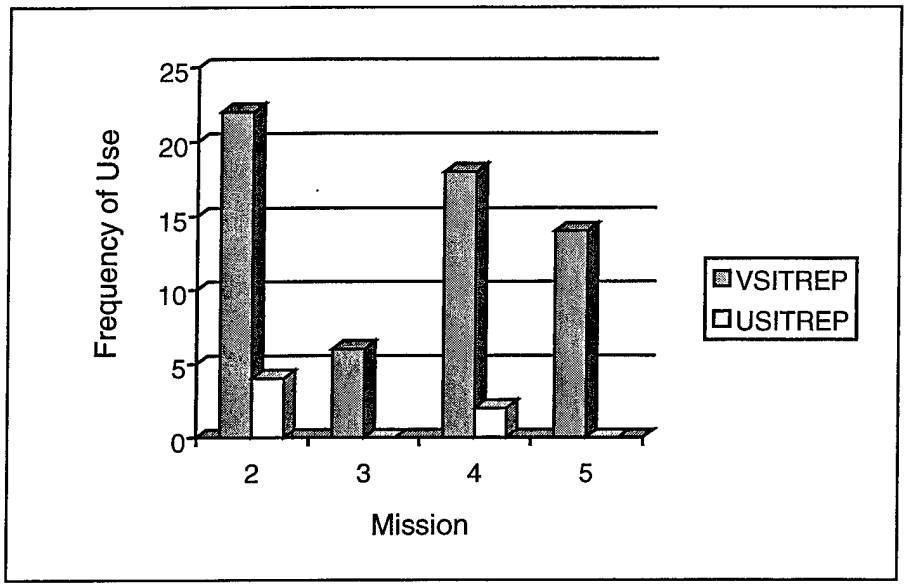


Figure B-1. Average total frequency of situation report use for commander and staff by mission.

2. Spot Report (SPOTREP) Use
   a. Operational definition: Calculate the frequency, type (individual vehicle or unit) and duration of use of the SPOTREP tool during the mission by node position. Additionally, record the frequency and duration of the "Auto On" feature. The data should be collected for each staff position.
   b. Rationale: Data may indicate the frequency of the staff's monitoring of activities and unit status during a mission. Measures the behaviors related to responding to requests for information, providing constructive suggestions, and obtaining information about the outcomes of decisions and actions where appropriate.
   c. Recommended output formats:
      1) Table – Frequency, type, and duration of SPOTREP use by mission by position.
      2) Graph – Frequency of SPOTREP use by mission by position (line graph).
      3) Graph – Average frequency of SPOTREP use by mission (bar graph).
   d. Data:
      1) Sample table and bar graphs provided in the main body of the report.
      2) Line graph not produced.

3. Commander's Critical Information Requirements (CCIR)
   a. Operational definition: Record the amount of time each node position uses and/or initiates use of the CCIR tool.
   b. Rationale: Provides insight into whether the use of the CCIR tool increases the overall level of situational awareness within the staff and allows them to focus on the areas

deemed important by the commander. Measures the behaviors related to responding to requests for information, providing constructive suggestions, obtaining information about the outcomes of decisions and actions where appropriate, and modifying activities or making new plans or decisions based on follow-up information.

c. Recommended output formats:
   1) Table – Frequency and duration of use of CCIR tool by mission by position.
   2) Graph – Average duration of use of CCIR tool by mission (bar graph).

d. Data:
   1) Sample table and graph provided below.

Table B-4

Frequency and Duration of Priority Information Request Tool Use for Staff Members by Mission

Position	Mission 2		Mission 3		Mission 4		Mission 5	
	Frequency	Duration	Frequency	Duration	Frequency	Duration	Frequency	Duration
WP6	1	0:04	0	0:00	0	0:00	0	0:00
WP62	0	0:00	0	0:00	6	6:45	0	0:00
WP52	0	0:00	0	0:00	0	0:00	1	1:16
WP53	1	2:00	2	0:30	3	1:25	0	0:00
WP83	0	0:00	0	0:00	2	5:10	4	7:01
WP99	3	3:26	0	0:00	2	1:49	1	1:15
WP92	0	0:00	0	0:00	3	9:43	4	8:11
WP93	1	4:11	0	0:00	0	0:00	0	0:00

Note. Duration times are total use in minutes and seconds.

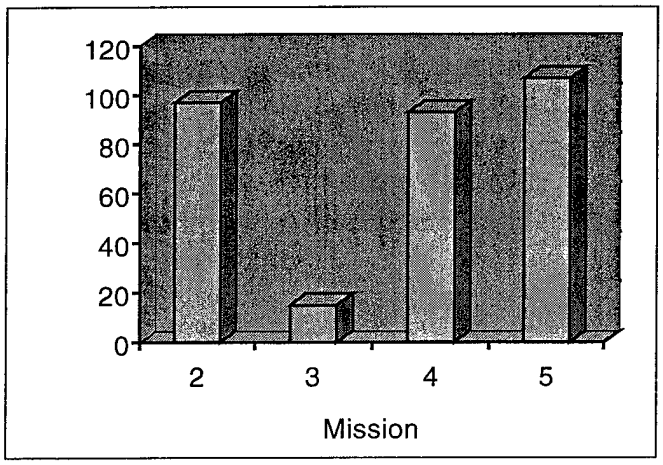


Figure B-2. Average duration of each Priority Information Request tool use for command and staff by mission.

4. Common Map Display
   a. Operational definition: At critical mission points, identify staff personnel who are displaying the same operations overlay and have their map set on the same center point and map scale. Additionally, identify staff personnel who are not displaying the same center point and map scale. Critical mission points are: first indirect fire engagement with OPFOR; first direct fire engagement with OPFOR; first friendly casualty; friendly losses exceed 30%; and last engagement during mission.
   b. Rationale: Differences in map center points and scales may indicate staff personnel are not monitoring significant activity during the mission. Measures the behaviors related to responding to requests for information, providing constructive suggestions, and obtaining information about the outcomes of decisions and actions where appropriate.
   c. Recommended output formats:
      1) Table – Staff personnel who are displaying the same operations overlay and have their map set on the same center point and map scale by position by critical mission points by mission.
      2) Graph – Number of staff personnel who are displaying the same operations overlay and have their map set on the same center point and map scale by critical mission points by mission (bar graph).
      3) Graph – Percentage of staff personnel who are displaying the same operations overlay and have their map set on the same center point and map scale by critical mission points by mission (bar graph).
   d. Data:
      1) This measure was not implemented during BCR IV.

5. Picture Consistency
   a. Operational definition: On a map or representation of a map, display those non-instrumented (those that don't automatically have their position and status reported by the SC^4 system) friendly higher, adjacent, and subordinate units and weapon systems that are not shown on the PVD. In addition, for those friendly force (higher, adjacent, and subordinate) non-instrumented units or weapon systems displayed on the PVD, calculate the time between the completion of a movement (greater than 200 meters) of the unit or system and the updating of its location on the PVD at fixed points in time.
   b. Rationale: The data may indicate the degree to which friendly force information not automatically displayed by the SC^4 system is shared among the staff to build situational awareness. Measures the behaviors related to responding to requests for information; providing construction suggestions, and obtaining information about the outcomes of decisions and actions where appropriate.
   c. Recommended output formats:
      1) Table – Time between completion of movement greater than 200m and updating of location on PVD of non-instrumented friendly systems by mission.
      2) Picture – Non-instrumented friendly higher, adjacent, and subordinate units not shown on PVD at fixed points in time.
      3) Picture – Displayed world (DW) vs. real world (RW) for non-instrumented friendly systems with use of different colors.
   d. Data:
      1) This measure was not implemented during BCR IV.

6. Operations Overlay Feedback
   a. Operational definition: For the 10 minutes following the opening of an operations overlay, identify staff personnel who are displaying the same operations overlay, have their map set on the same center point and map scale, and are at the same time talking on the radio with the creator of the overlay.
   b. Rationale: If staff personnel are viewing the same screen and are talking on the radio, they are probably discussing something on the screen. If this discussion immediately follows the opening of an overlay, then it will be assumed the creator of the overlay is receiving feedback. Measure the behaviors related to responding to requests for information, providing constructive suggestions, observing and tracking of team members' performance, and listening to other team member communications.
   c. Recommended output format:
      1) Table – Number of times a staff member and the creator of an operations overlay are displaying the created operations overlay, have their maps set on the same center point and map scale, and are talking at the same time on the radio; by operations overlay by mission.
   d. Data:
      1) This measure was not implemented during BCR IV.

Shared Situational Awareness

1. Map Area
   a. Operational definition: Based on the currently visible map area on a user's PVD/SC^4, calculate the square kilometers of the battlefield displayed. The view size may be affected by the size of the open PVD/SC^4 window, the scale of the map selected, and the use of SC^4 tools. The size should be calculated for each change in the map scale or resizing of the Modular Semi-Automated Forces (ModSAF) window and then averaged for the mission. Additionally, the center point of each resized PVD map should be computed and reported by grid. These data should be reported for each node position at fixed points in time (TBD). Picture formats of this output should enable visual comparison of visible map relative to total map and relative to other users.
   b. Rationale: The average amount of the battlefield that is displayed during a mission may indicate an appropriate and efficient use of the SC^4 system by the staff member. Measures the behaviors related to maintaining the big picture, identifying potential or anticipated problems, remaining aware of resources available, noting deviations from steady state, and recognizing need for action.
   c. Recommended output formats:
      1) Table – Average km^2 of battlefield displayed for each person for each mission or fixed points in time
      2) Table – Center point of PVD map for each person at fixed points in time
      3) Picture – Screen dumps of various peoples' PVD maps at same point in time
   d. Data:
      1) Sample table and pictures provided in the main body of the report.

2. Sensor Coverage
   a. Operational definition: On a two- or three-dimensional map or picture representation of a map depict the location of all enemy systems simulated at a specified time–corresponds to RW; depict the location of all enemy systems sensed at a specified time–corresponds to sensed world (SW); depict the location of all enemy systems displayed on the overall PVD/SC^4 map area–corresponds to DW; and depict the location of all enemy systems currently visible on a user's PVD/SC^4 at a specified time–corresponds to visible world (VW). Differentiate by color (or some other means) each of these enemy worlds and blue forces (BLUFOR) systems that have their location automatically displayed by the SC^4 system.
   b. Rationale: The display may indicate the degree to which each staff member or other personnel has situational awareness. Measures the behaviors related to maintaining the big picture, identifying potential or anticipated problems, remaining aware of resources available, noting deviations from steady state, and recognizing need for action.
   c. Recommended output format:
      1) Picture – At fixed point in time, depict location of all enemy systems:
         (a) currently visible on each user's PVD (VW) in red.
         (b) displayed on the overall PVD map area (DW) in faded red.
         (c) sensed (SW) in pale red.
         (d) simulated (RW) in pink.
   d. Data:
      1) This measure was not implemented during BCR IV.

3. Satellite Coverage
   a. Operational definition: On a map or representation of a map, display OPFOR military weapons systems that were detected solely by satellite intelligence gathering means, but not displayed on the PVD.
   b. Rationale: The data may indicate the degree to which intelligence information not automatically displayed by the SC^4 system is shared among the staff to build situational awareness. Measures the behaviors related to maintaining the big picture, identifying potential or anticipated problems, remaining aware of resources available, noting deviations from steady state, and recognizing need for action.
   c. Recommended output formats:
      1) Table – For systems detected only by satellite, number detected, time detected, time displayed on PVD by mission.
      2) Table – Number of systems detected by more than just satellite, number of systems not detected, number of systems detected but not displayed by mission.
      3) Picture – Display OPFOR military weapons systems detected only by satellite, both those displayed and not displayed on PVD in different colors.
   d. Data:
      1) This measure was not implemented during BCR IV.

4. Line of Sight (LOS)
   a. Operational definition: Record the amount of time each node position uses and/or initiates the LOS tool during the mission.

b. Rationale: Provides insight into whether the LOS tool assists the commander and planning staff in obtaining a better appreciation of the effects of terrain (OCOKA) in identifying and selecting axis of advance, battle positions, attacks by fire (ABFs), and supports by fire (SBFs) that support the scheme of maneuver. More precision in locating the best possible fighting locations, engagement areas, effects boxes and routes for inclusion in orders and plans may increase the probability that the mission can be accomplished at the least possible cost. Measures the behaviors related to maintaining the big picture, identifying potential or anticipated problems, remaining aware of resources available, noting deviations from steady state, and recognizing need for action.

c. Recommended output formats:
   1) Table – Frequency and duration of use of LOS tool by mission by position.
   2) Graph – Average duration of use of LOS tool by mission (bar graph).

d. Data:
   1) The Line of Sight measure was based on use of the Terrain Intervisibility tool. The results obtained by this measure were also used to explore the team process skill dimension of Adaptability (see Table 9 and Figure 10 in the main body of the report).

5. Surprise Attack

a. Operational definition: Tally the total number of flank or rear engagements on OPFOR and BLUFOR vehicles: attacks from a position that is greater than 45 degrees and less than 315 degrees of the hull orientation of the vehicle being attacked. The orientation of the vehicle is considered to be zero degrees for this calculation. If possible, express the total number of engagements, the number of flank or rear engagements and the number of engagements fitting this criteria divided by the total number of OPFOR or BLUFOR vehicles to provide a percentage of flank or rear engagements. Data collection should start at first direct fire engagement with OPFOR.

b. Rationale: Indicates situational awareness within battalion. Measures the behaviors related to maintaining the big picture, identifying potential or anticipated problems, remaining aware of resources available, noting deviations from steady state, and recognizing need for action.

c. Recommended output formats:
   1) Table – Frequency and percentage of flank or rear engagements by mission.
   2) Picture – For every flank or rear engagement in a mission, show location of BLUFOR in relation to OPFOR when they were attacked.

d. Data:
   1) Sample table provided below.
   2) Picture not produced.

Table B-5

Blue Forces (BLUFOR) and Opposing Forces (OPFOR) Engagements from Flank or Rear

	Mission 2	Mission 3	Mission 4	Mission 5
BLUFOR Engagements				
Total OPFOR Vehicles	402	582	570	391
# OPFOR Vehicles Engaged Flank/Rear	207	91	154	137
% of Rear/Flank Engagements	51	16	27	35
OPFOR Engagements				
Total BLUFOR Vehicles	318	287	366	294
# BLUFOR Vehicles Engaged Flank/Rear	58	23	145	67
% of Rear/Flank Engagements	18	8	40	23

6. Collateral Damage
   a. Operational definition: Report each instance of attack on BLUFOR non-instrumented vehicles and/or personnel by indirect non-line-of-sight (NLOS) weapons systems under battalion control during a mission. Data should reflect firing unit ID, echelon of controlling headquarters, type of weapon, time of engagement, and damage to the targeted BLUFOR unit. In addition, report each instance of attack on non-combatant non-instrumented vehicles and/or personnel by indirect NLOS weapons systems under battalion control during a mission. Data should reflect firing unit ID, echelon of controlling headquarters, type of weapon, time of engagement, and damage to the targeted non-combatant vehicle and/or personnel.
   b. Rationale: Data may indicate whether the staff is posting to the PVD the activities, location, and/or movement of non-instrumented BLUFOR units and of non-combatants within the battalion's area of responsibility. Measures the behaviors related to maintaining the big picture, identifying potential or anticipated problems, remaining aware of resources available, noting deviations from steady state, and recognizing need for action.
   c. Recommended output formats:
      1) Table – Firing unit ID, echelon of controlling headquarters, type of weapon, time of engagement, and type of damage for each instance of attack on BLUFOR non-instrumented vehicles and/or personnel by indirect NLOS weapons systems under battalion control by mission.
      2) Table – Firing unit ID, echelon of controlling headquarters, type of weapon, time of engagement, and type of damage for each instance of attack on non-combatant non-instrumented vehicles and/or personnel by indirect NLOS weapons systems under battalion control by mission.
      3) Picture – Snapshot of battlefield for every instance of attack on either BLUFOR or non-combatant non-instrumented vehicles and/or personnel.

d. Data:
   1) Sample table provided below.
   2) Picture not produced.

Table B-6

Incidents of Collateral Damage for Each Mission

Mission	Target	Firing Unit	Echelon	Type of Weapon	Time	Damage	Effect
2	Civilian_Car	_3M42	Plt	120Mor	13:52:09	Hit	KL
	Civilian_Car	_3M44	Plt	120Mor	14:17:31	Hit	
	Civilian_Car	_3M41	Plt	120Mor	14:41:20	Hit	
	Civilian_Car	_3M44	Plt	120Mor	14:41:28	Hit	
3	Civilian_DI	_3K15	Company	FV	15:31:40	Hit	KL
	Civilian_DI	_3K15	Company	FV	15:31:41	Hit	KL
	Civilian_DI	_3L34	Company	FV	15:34:24	Hit	KL
	Civilian_DI	_3L34	Company	FV	15:34:24	Hit	KL
4	None						
5	DI_Squad_Std	_3K16	Company	FV	15:12:22	Hit	
	DI_Squad_Std	_3K16	Company	FV	15:12:23	Hit	
	DI_Squad_Std	_3K16	Company	FV	15:16:19	Hit	
	DI_Squad_Std	_3K16	Company	FV	15:16:21	Hit	
	DI_Squad_Std	_3K16	Company	FV	15:21:05	Hit	
	DI_Squad_Std	_3K16	Company	FV	15:21:07	Hit	
	DI_Squad_Std	_3K16	Company	FV	15:23:07	Hit	
	DI_Squad_Std	_3K16	Company	FV	15:23:07	Hit	
	DI_Squad_Std	_3K16	Company	FV	15:27:28	Hit	
	DI_Squad_Std	_3K16	Company	FV	15:27:29	Hit	
	DI_Squad_Std	_3K16	Company	FV	15:28:39	Hit	
	DI_Squad_Std	_3K16	Company	FV	15:28:39	Hit	KL
	Civilian_Car	_3K23	Company	FV	17:10:34	Hit	KL
	Civilian_Car	_3K23	Company	FV	17:10:35	Hit	KL
	Civilian_Car	_3K23	Company	FV	17:11:42	Hit	KL
	Civilian_Car	_3K23	Company	FV	17:11:43	Hit	KL
	Civilian_Car	_3K23	Company	FV	17:11:51	Hit	KL
	Civilian_Car	_3K23	Company	FV	17:11:57	Hit	KL
	Civilian_DI	_3K23	Company	FV	17:11:58	Hit	KL
	Civilian_Car	_3K23	Company	FV	17:13:22	Hit	KL
	Civilian_SUV	_3M41	Plt	120Mor	18:39:16	Hit	

Note. 120Mor = 120mm Mortar Weapon System; FV = Fighting Vehicle; DI = dismounted infantry; KL = kill; SUV = sport utility vehicle.

7. SITREP Lag
   a. Operational definition: Calculate the difference in time between significant changes in friendly platoon status (e.g., change in overall unit status from green to amber, amber to red, red to black) and the initiation of a query by a staff member that notes that change.
   b. Rationale: A short time between a change in unit status and a query on the unit status by a staff member may indicate that the staff is closely monitoring the status of subordinate units and may have indicate the level of situational awareness within the staff. Measures the behaviors related to maintaining the big picture, identifying potential or anticipated problems, remaining aware of resources available, noting deviations from steady state, and recognizing need for action.
   c. Recommended output formats:
      1) Table – Time between changes in friendly platoon status (e.g., (G)reen, (A)mber, (R)ed, (B)lack) and initiation of staff member query by mission by platoon.
      2) Graph – Time platoon changed status vs. time queried by mission by platoon (bar graph).
   d. Data:
      1) This measure was not implemented during BCR IV.

Communication

1. Whiteboard Use
   a. Operational definition: Calculate the number of Whiteboard files residing on each staff member's workstation for each mission.
   b. Rationale: Data may indicate the extent to which the Whiteboard system is being used to coordinate activities within the staff. Measures the behaviors related to passing complete information to correct members and providing information in complete, accurate, timely, and efficient manner.
   c. Recommended output formats:
      1) Table – Number of Whiteboard files residing on each staff member's workstation by mission by position.
      2) Graph – Average number of Whiteboard files residing on the workstations by mission by node (bar graph).
   d. Data:
      1) Sample table and graph provided in the main body of the report.

2. Radio Communications Pattern
   a. Operational definition: Calculate the use of battalion command and operations-intel radio nets and Whiteboard conferencing by battalion staff members at critical points during a mission. Critical points are: a first indirect fire engagement with OPFOR; first direct fire engagement with OPFOR; first friendly casualty; friendly losses exceed 30%; and last engagement during mission. The critical event time period to be measured are from the initiation of the event and for 10 minutes thereafter. Compare the time usage during the critical events with the average usage of the Whiteboard and radio during the mission.
   b. Rationale: Analysis of the communication patterns will determine if battalion staff personnel are communicating more with each other during critical mission events.

Measures the behaviors related to passing complete information to correct members and providing information in complete, accurate, timely, and efficient manner.

c. Recommended output formats:
   1) Table – Number of times and average time of each usage of radio and whiteboard during critical points and number of times and average time of each usage of radio and whiteboard at other times by mission.
   2) Graph – Percentage of total of number of times radio used for each critical point during a mission by mission (bar graph).
   3) Graph – Percentage of total of number of times whiteboard used for each critical point during a mission by mission (bar graph).

d. Data:
   1) Sample tables provided below. The data are reported for the 10-minute period following each critical event to ensure that the commander and primary staff had an opportunity to recognize that a critical event had occurred and to react to it.
   2) Graphs not produced.

Table B-7

Radio Communications Patterns for the Senior Staff Members during Mission 2

Mission 2 – Defend in Sector. STARTEX 1235 hours.				
	WP6	WP5	WP88	WP99
Critical Event: First Indirect Fire Engagement (1330 – 1340 hours)				
Total number of transmissions	0	21	20	0
Total transmission time	0	112.3	142.2	0
Average transmission time	0	5.3	7.1	0
Critical Event: First Direct Fire Engagement (1335 - 1345 hours)				
Total number of transmissions	4	14	20	0
Total transmission time	12.5	86.3	96.5	0
Average transmission time	3.1	6.2	4.8	0
Critical Event: Last Engagement (1935 – 1945 hours)				
Total number of transmissions	5	2	14	12
Total transmission time	46.4	5.2	82.1	46.2
Average transmission time	9.3	2.6	5.9	3.9
Radio Transmissions During Non-Critical Event Periods				
Total number of transmissions	352	572	119	65
Total transmission time	2372.1	4218.5	425.9	321.7
Average number of transmissions[a]	1	2	1	1
Average transmission time	6.7	7.4	3.6	4.9

Note. STARTEX = start of exercise.
[a] Data averaged for a 10-minute block during non-critical event periods for entire mission.

Table B-8

Radio Communications Patterns for the Senior Staff Members during Mission 4

Mission 4 – Defend in Sector. STARTEX 1230 hours.				
	WP6	WP5	WP88	WP99
Critical Event: First Indirect Fire Engagement (1516 –1526 hours)				
Number of transmissions	5	5	12	0
Total transmission time	14.1	44.1	176.6	0
Average transmission time	2.8	8.8	14.7	0
Critical Event: First Direct Fire Engagement (1546 - 1556 hours)				
Number of transmissions	5	4	11	0
Total transmission time	25.8	23.1	303.0	0
Average transmission time	5.2	5.8	27.5	0
Critical Event: Last Engagement (1919 - 1929 hours)				
Number of transmissions	9	0	0	0
Total transmission time	91.1	0	0	0
Average transmission time	10.1	0	0	0
Radio Transmissions During the Rest of the Exercises (ENDEX at 1929)				
Number of transmissions	242	401	233	58
Total transmission time	1526.3	2811.2	1684.2	224.7
Average transmission time	6.3	7.0	7.2	3.9

Note. ENDEX = end of exercise; STARTEX = start of exercise.

Table B-9

Radio Communications Patterns for the Senior Staff Members during Mission 5

Mission 5 – Move to Engage. STARTEX 1240 hours.				
	WP6	WP5	WP88	WP99
Critical Event: First Indirect Fire Engagement (1245 –1255 hours)				
Number of transmissions	0	16	9	3
Total transmission time	0	114.1	74.7	18.7
Average transmission time	0	7.1	8.3	6.2
Critical Event: First Direct Fire Engagement (1410 - 1420 hours).				
Number of transmissions	0	24	2	0
Total transmission time	0	180.1	11.5	0
Average transmission time	0	7.5	5.7	0
Critical Event: Last Engagement (1849 – 1859 hours).				
Number of transmissions	0	10	0	0
Total transmission time	0	41.1	0	0
Average transmission time	0	4.1	0	0
Radio Transmissions During the Rest of the Exercises (ENDEX at 1900)				
Number of transmissions	0	456	92	34
Total transmission time	0	3306.4	751.7	179.9
Average transmission time	0	7.3	8.2	5.3

Note. ENDEX = end of exercise; STARTEX = start of exercise.

3. Personnel Initiating Whiteboard Conferences
   a. Operational definition: Record by staff position, the number of Whiteboard conferences, lasting 3 minutes or more, initiated during the mission.
   b. Rationale: The number of staff personnel initiating Whiteboard conferences, and their frequency of initiation may indicate the level of communication within the battalion staff. Measures the behaviors related to passing complete information to correct members and providing information in complete, accurate, timely, and efficient manner.
   c. Recommended output format:
      1) Table – Number of whiteboard conferences lasting 3 minutes or more by staff position by mission.
   d. Data:
      1) Sample table provided below.

Table B-10

Staff Personnel Initiating Whiteboard Conferences Lasting Three Minutes or More

Start	Stop	Duration	From	To
Mission 4 – Defend in Sector				
18:31:18	18:34:59	00:03:41	WP5	Higher Headquarters
Mission 5 – Move to Engage				
13:32:36	13:53:40	00:21:04	WP5	Higher Headquarters
14:17:15	14:32:43	00:15:28	WP5	Company Commander
14:18:10	14:32:33	00:14:23	WP88	WP5
14:18:20	14:32:34	00:14:14	WP99	WP5
18:48:10	18:51:25	00:03:15	WP52	Higher Headquarters

4. Information Flow
   a. Operational definition: Direction and volume of communication methods (e.g., Whiteboard, e-mail, Alert tool, PVD clearboard) between the commander and staff.
   b. Rationale: May indicate which SC^4 components were most useful between the commander and staff. Could also assist in further SC^4 system design and development. Measures the behaviors related to passing complete information to correct members and providing information in complete, accurate, timely, and efficient manner.
   c. Recommended output format:
      1) Picture – Direction of flow of information
   d. Data:
      1) This measure was not implemented during BCR IV.

Coordination

1. Overlay Use
   a. Operational definition: Determine the number of staff members that are showing, on their PVD, the same operations overlay file that the battalion commander is showing on his PVD at fixed points in time (TBD). These data should be reported as a percentage by dividing the number of staff members showing the same overlay file as the commander by the total number of staff members. If possible, report by duty position and operations overlay file name, those staff members who are not showing the same file as the commander at fixed points in time.
   b. Rationale: The data may indicate whether the commander and his staff are using the same graphic control measures to monitor and control subordinate units. If they are not, then there is a significant potential for miscommunication and a breakdown of situational awareness within the unit. Measures the behaviors related to synchronizing actions, passing information in timely and efficient manner, and facilitating performance of other team members.

c. Recommended output formats:
   1) Table – Frequency and percentage of staff members showing on their PVD the same operations overlay file as the Battalion Commander at fixed points in time by mission by overlay file
   2) Graph – Average number of staff members showing on their PVD the same operations overlay file as the Battalion Commander by fixed points in time by mission

d. Data:
   1) Sample table provided below.
   2) Graph not produced.:

Table B-11

Frequency and Percent of Staff Members Displaying Same Overlay Files as Commander by Mission

Mission and Files	Position											Frequency	Percent[a]
	WP	5	52	53	62	63	82	83	88	92	93	99		
Mission 2														
Overlay 2.1			✓			✓	✓	✓	✓	✓	✓	✓	8	62%
Overlay 2.2			✓					✓	✓	✓		✓	5	38%
Overlay 2.3		✓	✓	✓		✓	✓	✓	✓	✓	✓	✓	10	77%
Overlay 2.4		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	11	85%
Overlay 2.5		✓	✓	✓	✓	✓	✓	✓	✓		✓	✓	10	77%
Mission 3														
Overlay 3.1		✓	✓	✓	✓					✓			5	38%
Overlay 3.2			✓	✓			✓	✓	✓		✓		6	46%
Overlay 3.3			✓	✓		✓	✓	✓	✓	✓	✓	✓	9	69%
Overlay 3.4			✓	✓		✓	✓	✓	✓	✓		✓	8	62%
Mission 4														
Overlay 4.1						✓		✓		✓	✓		4	31%
Overlay 4.2					✓	✓	✓		✓				4	31%
Overlay 4.3		✓											1	8%
Overlay 4.4		✓	✓		✓	✓	✓	✓	✓	✓	✓	✓	10	77%
Overlay 4.5				✓	✓			✓		✓	✓		5	38%
Overlay 4.6				✓	✓			✓		✓	✓		5	38%
Overlay 4.7		✓	✓			✓			✓	✓		✓	6	46%
Overlay 4.8			✓		✓		✓	✓	✓	✓		✓	7	54%
Overlay 4.9		✓	✓					✓	✓	✓	✓	✓	7	54%
Overlay 4.10		✓	✓	✓	✓	✓	✓		✓	✓		✓	9	69%
Overlay 4.11		✓	✓	✓	✓	✓	✓		✓	✓		✓	9	69%
Mission 5														
Overlay 5.1													0	0%
Overlay 5.2		✓	✓	✓			✓	✓	✓	✓	✓	✓	9	69%
Overlay 5.3		✓	✓	✓			✓	✓	✓	✓	✓	✓	9	69%

[a] Percentages are based upon 11 staff members. The two Sensor Non-commissioned Officers (WP84 and WP94) did not have SC^4 systems to exchange overlay files.

2. Whiteboard Commonality
   a. Operational definition: Determine the number of staff members that have, within their Whiteboard directory, the same Whiteboard files that the commander and deputy commander have at fixed points in time (TBD). These data should be reported as a percentage by dividing the number of staff members showing the same Whiteboard files as the commander and deputy commander by the total number of staff members. Report by duty position, those staff members who are not showing the same Whiteboard files as the commander and deputy commander at fixed points in time. In addition, determine the number of staff members that have within their Whiteboard system, the same files that the node officer in charge (OIC) has at fixed points in time (TBD). These data should be reported as a percentage by dividing the number of staff members having the same Whiteboard files as the node OIC by the total number of staff members in the node. Report by duty position, those staff members who are not showing the same Whiteboard files as the node OIC at fixed points in time.
   b. Rationale: In previous experimentation, the unit operations orders (OPORDs) were transmitted to the staff and subordinate units through the use of Whiteboard files and through the use of PVD overlay files. While the content of the Whiteboard files can not be automatically derived, if the unit uses Whiteboard files to transmit OPORDs, the data may indicate whether the commander and his staff are using the same OPORD to monitor and control subordinate units. The initial comparison between the commander and deputy commander is a method to ensure that the OPORD file or files can be identified without having to analyze the content of the file. If the commander and staff are not sharing the same OPORD, there is a significant potential for miscommunication and a breakdown of situational awareness within the unit. Measures the behaviors related to synchronizing actions, passing information in timely and efficient manner, and facilitating performance of other team members.
   c. Recommended output formats:
      1) Table – Frequency and percentage of staff members showing in their whiteboard directories the same whiteboard files as the Battalion Commander and Deputy Commander at fixed points in time by mission by whiteboard file.
      2) Table – Frequency and percentage of staff members showing in their whiteboard directories the same whiteboard files as their node OIC at fixed points in time by mission by whiteboard file.
      3) Picture – List of all whiteboard files on Battalion Commander's system, and highlight all the ones that everyone else also has.
   d. Data:
      1) Sample table provided below. Since the squadron commander did not have any Whiteboard files, the measure is reported with the deputy commander as the comparison point. The Whiteboard file selected for each mission was the squadron OPORD.
      2) Picture not produced.

Table B-12

Whiteboard Files Staff Members Had in Common with the Deputy Commander (WP5)

Position	Whiteboard Files			
	Mission 2	Mission 3	Mission 4	Mission 5
WP6				
WP62		✓	✓	✓
WP63		✓	✓	✓
WP52		✓	✓	✓
WP53		✓	✓	
WP88	✓	✓	✓	✓
WP82			✓	✓
WP83		✓	✓	✓
WP84		✓		✓
WP99		✓	✓	✓
WP92			✓	✓
WP93			✓	✓
WP94				
Iron 6		✓	✓	✓
Killer 6		✓	✓	✓
Lightning 6		✓	✓	✓
Mad Dog 6		✓	✓	✓
Total	2	13	15	15
Percent	11%	72%	83%	83%

3. Targeting
   a. Operational definition: Calculate, for each request for indirect fire initiated within the staff, whether the requester performed a SPOTREP query on the target identified in the fire request immediately before transmitting the request for fire. The data should be reported by each node position. In addition, calculate, for each request for indirect fire received by the staff officer responsible for approving indirect fire missions, whether he performed a SPOTREP query on the target identified in the fire request immediately before approving the request for fire. The data should be reported by each node position identified as approval authority for indirect fire support requests.
   b. Rationale: The SC^4 system display of OPFOR positions may not provide the most current position if the OPFOR was moving and wasn't being continuously tracked by a friendly sensor. As a result, the request for fire may not be accurate. The data may indicate whether the requester and/or the fire mission coordinator understand the capabilities and limitations of the SC^4 system to display OPFOR data. Measures the

behaviors related to synchronizing actions, passing information in timely and efficient manner, and facilitating performance of other team members.

c. Recommended output formats:
   1) Table – For each request for indirect fire, frequency and time of SPOTREP query on target before transmitting request and time of transmittal by mission by position.
   2) Table – For each request for indirect fire received, time of receipt of request, frequency and time of SPOTREP query on target before approving request, and time of approval by mission by approval authority.
   3) Graph – Number of kills per indirect fire request when: SPOTREP query performed before request, SPOTREP query performed before approval, SPOTREP queries. performed before request and approval, and SPOTREP queries not performed (line graph).

d. Data:
   1) Sample table provided below.
   2) Graph not produced.

Table B-13

Use of Spot Report (SPOTREP) for Target Verification Before Initiating Fire Mission Request

Mission	Action	Position							
		WP 53	63	82	83	84	92	93	94
	# Fire Mission Requested	1	95	1	5	12	15	5	1
2	# Fire Mission w/other Adjacent Targets	1	68	--	--	6	--	--	--
	# SPOTREP/Fire Mission Pairings	--	9	--	--	--	--	--	--
	# Fire Mission Requested	--	57	--	--	--	--	--	--
3	# Fire Mission w/other Adjacent Targets	--	53	--	--	--	--	--	--
	# SPOTREP/Fire Mission Pairings	--	6	--	--	--	--	--	--
	# Fire Mission Requested	--	141	--	--	--	--	--	--
4	# Fire Mission w/other Adjacent Targets	--	122	--	--	--	--	--	--
	# SPOTREP/Fire Mission Pairings	--	4	--	--	--	--	--	--

4. Fire Support Coordination
   a. Operational definition: Calculate the ratio of OPFOR kills due to indirect fire from units controlled by the battalion staff to OPFOR kills due to direct fire controlled by battalion subordinate units. Further breakdown the data to show the percentage that each type of indirect fire system contributed to the total number of OPFOR kills.
   b. Rationale: The data may indicate that the staff has sufficient situational awareness and the ability to synchronize the fire support assets under their control to inflict significant damage to the OPFOR before the OPFOR can close to within direct fire range. The further away from the unit the OPFOR can be defeated the less risk there is to the unit. Measures the behaviors related to synchronizing actions, passing information in timely and efficient manner, and facilitating performance of other team members.

c. Recommended output formats:
   1) Table – Frequency and percentage of OPFOR kills due to indirect fire from units controlled by the battalion staff and frequency and percentage of OPFOR kills due to direct fire controlled by battalion subordinate units by mission.
   2) Graph – Frequency of OPFOR kills due to indirect fire and frequency of OPFOR kills due to direct fire by mission (bar graph).
   3) Table or Bar Graph – Frequency and percentage of OPFOR kills by each type of indirect fire system.
   4) Picture – At end of mission, OPFOR locations on battlefield (indirect kills in gray, direct kills in black, alive in red) in relation to BLUFOR locations.

d. Data:
   1) Sample table and graph provided in the main body of the report.
   2) Picture not produced.

5. Fire Engagements

a. Operational definition: Calculate the average range of OPFOR (ATGM, TERM, tank main gun, 35mm sabot, etc.) and BLUFOR (NLOS, PGMM, 35mm sabot, CKEM) weapon system engagements, by type against BLUFOR or OPFOR vehicles that were killed during a mission. For each OPFOR target engaged, calculate the total number of engagements against it by BLUFOR weapons systems from the first engagement until end of mission. Data collected should indicate the firing unit, echelon of its controlling headquarters, the type of weapon used, the engagement time, and should indicate at which time the OPFOR target was killed (firepower, mobility, and/or catastrophic).

b. Rationale: Killing OPFOR vehicles at or near the maximum effective range of a weapon may indicate efficiency and precision in execution. Killing BLUFOR vehicles at or near the maximum effective range of an OPFOR weapon may indicate lack of situational awareness and precision in execution. The data on multiple engagements of an OPFOR target may indicate whether the situational awareness provided by the SC^4 system and an effective unit fire coordination and distribution system reduced or prevented the needless expenditure of ammunition against targets already destroyed. Measures the behaviors related to synchronizing actions, passing information in timely and efficient manner, and facilitating performance of other team members.

c. Recommended output formats:
   1) Table – Average range of each OPFOR weapon system engagement against BLUFOR vehicles killed by mission by type of weapon.
   2) Table – Average range of each BLUFOR weapon system engagement against OPFOR vehicles killed by mission by type of weapon.
   3) Table – For each OPFOR target engaged more than once, total number of engagements against it by BLUFOR weapons systems by firing unit by echelon of its controlling headquarters, by type of weapon used, by engagement time, by time target killed, by mission.
   4) Picture – At end of mission, OPFOR kills engaged by multiple friendly weapons systems in black and OPFOR kills engaged only once in gray by mission.

d. Data:
   1) Sample table provided below.
   2) Picture not produced.

Table B-14

Average Opposing Forces (OPFOR) and Blue Forces (BLUFOR) Weapon System Engagement Ranges in Meters

Vehicle and Ammunition Type	Maximum Effective Range	Mission 2	Mission 3	Mission 4	Mission 5
OPFOR Weapon Systems					
T-90 (125mm)	2000	1060	--	2055	3160
T-90 (AT8)	4000	--	--	1342	--
BMP2 (30mm)	1000	--	--	246	--
BMP2 (AT5)	4000	--	--	1832	--
BMP3 (100mm)	4000	--	--	--	--
BMP3 (AT10)	4000	--	--	--	--
BTR-80	2000	294	2482	1939	--
BRDM2K	4000	--	2900	--	--
BM22[a]	140000	21899	34417	51392	59337
2S19	24700	18883	--	--	--
BLUFOR Weapon Systems					
FV (35mm)	2000	364	--	1406	--
FV (CKEM)	6000	764	--	1007	3354
FV (HG MP)	15000	7359	2543	7756	6241
FV (HG KE)	4000	1232	--	1244	--
FV (HG SAD)	25000	11405	18684	20782	21540
NLOS - CMM	15000	11960	11530	9093	11121
Javelin	4500	2826	886	2488	--
MAV APC - 25mm	2800	295	1628	866	1824
HIMARS[a]	140000	40281	31480	83584	55270
MAV Mortar (PGMM)	15000	10061	11276	10772	7125
MAV Mortar (HE)	7000	--	--	--	--

[a] Includes tactical missile launcher capability.

6. OPFOR Destruction
   a. Operational definition: Calculate the time from the first OPFOR engagement until OPFOR vehicle losses exceeded 70%. In addition, calculate the rate at which the OPFOR was killed in five minute intervals from the first engagement until the last OPFOR kill during the mission. The data collected should reflect cumulative numbers and results for each weapon type.
   b. Rationale: Rapid destruction of the OPFOR reduces risk of losses to the friendly unit and may indicate improvements in battle command capabilities over current capabilities. The rate of destruction may indicate the battle tempo during the mission and provide insight into the maximum capability of an SC^4 equipped unit to synchronize combat and combat support assets to efficiently destroy the OPFOR. Measures the behaviors related to synchronizing actions, passing information in timely and efficient manner, and facilitating performance of other team members.
   c. Recommended output formats:
      1) Table – In 5-minute intervals, the cumulative rate of OPFOR losses from first engagement until last OPFOR kill by mission by weapon type. Also include time from first OPFOR engagement until losses exceed 70%.
      2) Graph – In 5-minute intervals, average rate of OPFOR losses by mission (line graph).
      3) Picture – Snapshot of battlefield at first engagement and another when losses exceed 70%.
   d. Data:
      1) Sample table and graph provided in the main body of the report.
      2) Picture not produced.

Decision-Making

1. Unmanned Aerial Vehicle (UAV) Effectiveness
   a. Operational definition: For each UAV launch, calculate the % of OPFOR vehicles (by type [tank, infantry fighting vehicle, armored personnel carrier, artillery, air defense artillery, etc.]) that are first detected by the UAV under battalion control. A possible measurement method is the number of OPFOR vehicles first detected by UAV missions divided by the total number of OPFOR vehicles detected by all other sensors and weapon systems. The UAV sensor capabilities will be determined by the parameters established for the system prior to the start of the experiment. The UAV flight path will be determined by experimental unit personnel.
   b. Rationale: The SC^4 system enables the commander to visualize the battlefield more effectively. This should include the ability to visualize the capability of terrain to support combat unit maneuver and appreciate regions of the battlefield where information is incomplete or non-existent. If they are being used effectively, UAVs should be sent to those areas of interest that are not already being covered by other sensors. If they are sent to areas of interest already covered by sensors, then a case can be made that either the SC^4 system is not helping the commander better visualize the battlefield or the commander does not trust the situational awareness depicted by the SC^4 system. Measures the behaviors related to using all relevant available information and making decisions in a timely manner.

c. Recommended output formats:
   1) Table – Percent and number of OPFOR vehicles by type first detected by UAVs (frequency table) instead of by all other sensors and weapons systems.
   2) Picture – Flight paths of UAVs for each mission, as well as flight paths of micro-UAVs (MUAVs), etc., in different colors to see if UAVs were being used to best advantage.

d. Data:
   1) Sample table and picture provided in the main body of the report.

2. Length of Battalion Decision-Making Cycle (Operations Order)
   a. Operational definition. Calculate the length of time from the receipt of a higher headquarters Whiteboard file with OPORD in the subject line until the battalion distributes a PVD operations overlay file to all PVD users in the battalion.
   b. Rationale: The shorter the time interval from receipt of the higher headquarters OPORD until the battalion issues its OPORD to its subordinates, the more time the subordinates will have to prepare for the upcoming mission. Increased preparation time should result in a more favorable outcome (attainment of the mission objective, reduced personnel casualties and material losses, conservation of ammunition and fuel, etc.). Measures the behaviors related to using all relevant available information, making mostly correct decisions based on information available, and making decisions in a timely manner.
   c. Recommended output format:
      1) Graph – Length of time from receipt of higher headquarters Whiteboard file with OPORD in the subject line until the battalion distributes a PVD operations overlay file to all PVD users in the battalion by mission (bar graph).
   d. Data:
      1) This measure was not implemented during BCR IV.

3. Length of Battalion Decision-Making Cycle (Platoon Movement)
   a. Operational definition. Calculate the length of time from the receipt of a higher headquarters Whiteboard file with OPORD in the subject line until at least one platoon leader in three different companies simultaneously shift their locations in excess of 2000 meters toward the OPFOR or away from the OPFOR.
   b. Rationale: The shorter the time interval from receipt of the higher headquarters OPORD until the battalion issues its OPORD to its subordinates, the more time the subordinates will have to prepare for the upcoming mission. Increased preparation time should result in a more favorable outcome (attainment of the mission objective, reduced personnel casualties and material losses, conservation of ammunition and fuel, etc.). Measures the behaviors related to using all relevant available information, making mostly correct decisions based on information available, and making decisions in a timely manner.
   c. Recommended output format:
      1) Graph – Length of time from receipt of higher headquarters Whiteboard file with OPORD in the subject line until at least one platoon leader in three different companies simultaneously shift their locations in excess of 2000 meters toward the OPFOR or away from the OPFOR (bar graph).
   d. Data:
      1) This measure was not implemented during BCR IV.

4. Information Retrieval by the Commander
   a. Operational definition: Calculate the amount of information the commander retrieves on his own (e.g., SITREPs) that other staff normally retrieve for him.
   b. Rationale: May indicate that the SC^4 system is providing sufficient situational awareness to the commander to compensate for a reduced battle staff. Measures the behaviors related to using all relevant available information and making mostly correct decisions based on information available.
   c. Recommended output formats:
      1) Table – Frequency and duration of SC^4 system tool use.
   d. Data:
      1) This is a composite measure which used the data from the Terrain Analysis, SITREP Use, SPOTREP Use, and PIR measures as indicators of the amount of information the commander accessed on his own beyond that displayed automatically on his PVD. During BCR IV, the data obtained by the cited measures indicate that the commander did not query the SC^4 system for additional information.