

Research Note 80-6

LEVEL

12

ADA 082707

DEVELOPMENT OF A METHODOLOGY FOR MEASURING TRANSFER
OF TRAINING EFFECTS FOR TACTICAL TRAINING SYSTEMS

Howard G. Hill and Gary Kress
Human Resources Research Organization

DTIC
ELECTE
APR 4 1980
C

ARI FIELD UNIT, USAREUR



U. S. Army

Research Institute for the Behavioral and Social Sciences

September 1979

Approved for public release; distribution unlimited.

DDC FILE COPY

80 4 4 045

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

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

JOSEPH ZEIDNER
Technical Director

FRANKLIN A. HART
Colonel, US Army
Commander

Research accomplished
for the Department of the Army
Human Resources Research Organization

NOTICES

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

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

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

18 APR 19 20 20-1

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER Research Note 80-6	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Development of a Methodology for Measuring Transfer of Training Effects for Tactical Training Systems	5. TYPE OF REPORT & PERIOD COVERED Final Report, October 1978 - September 1979	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Howard G. Hill Gary/Kress	8. CONTRACT OR GRANT NUMBER(s) MDA 903-78-C-2042	
9. PERFORMING ORGANIZATION NAME AND ADDRESS Human Resources Research Organization (HumRRO) 300 North Washington Street Alexandria, Virginia 22314	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS Army Project 62722A764 Task G, Work Unit 1	
11. CONTROLLING OFFICE NAME AND ADDRESS U.S. Army Research Institute for the Behavioral and Social Sciences PER-OI 5001 Eisenhower Ave., Alex., Va. 22333	12. REPORT DATE September 1979	
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) U.S. Army Research Institute for the Behavioral and Social Sciences, USAREUR Field Unit ODCSPER, HQ, USAREUR APO New York 09403	13. NUMBER OF PAGES 26 (plus appendices)	15. SECURITY CLASS. (of this report) Unclassified 137
16. DISTRIBUTION STATEMENT (of this Report) Approved for public release; distribution unlimited.		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report) --		
18. SUPPLEMENTARY NOTES --		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) Transfer of training, evaluation methodology, performance evaluation, tactical offensive performance indicators (Company/Team)		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) A methodology for the measurement of transfer of training between a battle simulation (Dunn-Kempf) and a tactical field exercise (REALTRAIN) was de- veloped. The methodology is based on evaluation of tactical performance in company level combined arms exercises (offensive mission) in both training environments. This report documents the development, tryout, and refinement of a tactical performance evaluation instrument and discusses the development and use of an evaluator training program.		

DD FORM 1473

EDITION OF 1 NOV 65 IS OBSOLETE

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

402260

JOB

20.

→ Attached to the report is research plan for determining the reliability of measurement for the instrument developed during this study and for conducting a transfer of training experiment between Dunn-Kempf training exercises and REALTRAIN field exercises.

The report and the performance evaluation instrument have application for evaluating the effectiveness of tactical training for Company/Team leaders.

Unclassified

FOREWORD

This report explains the development of a tactical performance evaluation instrument and a methodology for measuring transfer of training between a battle simulation and a tactical field exercise. The impetus for the developmental effort stems in part from the Army's need to determine the training value and appropriate role of battle simulations in the overall Army training repertoire. The current methodological effort represents a continuation of earlier Army Research Institute (USAREUR) research which evaluated learning within the Dunn-Kempf battle simulation. Actual conduct of a transfer of training experiment will be the objective of the next phase of the effort. The research reported here relates to the larger area of research dealing with evaluation of tactical performance and with reliability and comparability of performance measurement among different exercises, participating units, and performance evaluators.

The research plan attached to this report outlines the procedures for analyzing the reliability of measurement using the prescribed methodology and for conducting a transfer of training experiment.

The work reported here was performed at the Heidelberg office of the Human Resources Research Organization (HumRRO), under Contract No. MDA903-78-C-2042 with the U.S. Army Research Institute for the Behavioral and Social Sciences (ARI). Dr. William W. Haythorn was the Contracting Officer's Technical Representative.

The authors wish to express special appreciation to Cpt. Scott, Cpt. Winchell, Lt. Wann, Lt. Brislawn, Lt. Lott, Lt. Kluge and Sgt. Closs of the 8th Infantry Division for their cooperation and expertise provided during field tests of the evaluation instrument at Baumholder. Appreciation is also expressed to Maj. Bickel of the 3rd Infantry Division for coordinating data collection support during field exercises at Hohenfels. And finally gratitude is expressed for the technical input and insights afforded by Maj. Jim Kurtz of the Force Modernization Division of DCSOPS, HQ, USAREUR.

Accession For	
NTIS GRA&I	☒
DOC TAB	☐
Unannounced	☐
Justification	☐
By _____	
Distribution/	
Availability Codes	
Dist	Avail and/or special
A	

Development of a Methodology for Measuring Transfer of Training Effects for Tactical Training Systems

BRIEF

Requirement:

To develop a methodology to measure transfer of training from a battle simulation (Dunn-Kempf) to a tactical field exercise (REALTRAIN).

Procedure:

An instrument was developed for evaluating the tactical performance of Company/Team leaders (company commander, platoon leaders, FIST chief) in offensive combined arms exercises in either Dunn-Kempf or REALTRAIN tactical training environments. Evaluation of performance in the tactical exercises relies on the application of performance measures by qualified military evaluators. Tactical skills to be evaluated were identified from performance measures developed for earlier Dunn-Kempf research* and from analysis of tactical doctrine for the offensive missions.

Measuring transfer of training would require the evaluation and comparison, in REALTRAIN exercises, of experimental units that have participated in Dunn-Kempf training and control units that have not had Dunn-Kempf training. Reliability of performance measurement between exercises is critical to the interpretation of transfer effects. The approach taken to achieve reliability was through standardization of measurement procedures. The approaches to standardizing application of the measures were: (a) to delineate the components of tactical skills involved in each measure so that ratings would be made on the same bases; (b) to explicitly define the anchor points on the rating scale used; and (c) to develop an evaluator training program to explain the use of the measures. The performance evaluation instrument and evaluator training program were used and revised during two periods of field exercises.

As the final step in the current phase of the research effort, a two-stage research plan was developed (1) for determining the reliability of measurement of the evaluation methodology developed and (2) for conducting a transfer of training experiment.

Utilization:

The methodology developed was designed for experimental measurement of transfer of training effects between tactical training systems, but has the more general applicability of evaluating the effectiveness of (offensive) tactical training for Company/Team leaders. The evaluation methods and measures developed could further serve as a model for development of performance evaluation instruments for other tactical missions.

*Elmo E. Miller and J.M. Bachta, "An Experimental Evaluation of a Tactical Game for Company Level Training," Army Research Institute Draft Technical Report, January 1978.

TABLE OF CONTENTS

	<u>Page</u>
INTRODUCTION	1
Background	1
Measuring Transfer of Training	2
PURPOSE	4
METHOD	4
1. Develop Approach	4
2. Conduct Preliminary Observations	7
3. Develop Instrument	9
4. Develop Evaluator Training Program	20
5. Investigate Reliability in Field	22
6. Analyze Results	22
7. Develop a Transfer of Training Research Plan	25
CONCLUSIONS	26
APPENDIX A: Dunn-Kempf/REALTRAIN Tactical Performance Evaluation Form	A-1
APPENDIX B: Defensive Rating Scale for Previous Dunn-Kempf Study	B-1
APPENDIX C: Evolution of Original Dunn-Kempf Measures to Finalized Performance Evaluation Instrument	C-1
ATTACHMENT: Research Plan for Measuring Transfer of Training Between Tactical Training Systems	D-1

INTRODUCTION

The research reported in this study is concerned with the measurement of training transfer as it applies to military tactical training systems. Transfer of training refers to the phenomenon through which skills learned in one type of training setting affect performance when a similar situation is encountered in another setting. Because all tactical training is conducted in a non-combat environment, it relies on the assumption that tactical skills and knowledge will transfer to a combat situation when needed. This study dealt specifically with the problem of how to measure transfer of training from a board game battle simulation to a field tactical training environment.

BACKGROUND

In conducting tactical field training exercises at company and higher levels, the Army is often faced with problems relating to: accessibility of terrain of adequate dimensions; operating with understrength or fragmented units; high costs of POL and pyrotechnics for large-scale exercises; and, complexity of planning and coordinating support. In USAREUR the limited access to adequate maneuver areas for tactical training is a particularly severe and chronic problem.

Because of these problems, the U.S. Army has developed a number of battle simulation and wargaming techniques. These techniques range in complexity from small playing boards with manually controlled playing pieces involving a few players, to large-scale computer-controlled exercises for battalion and higher levels of play. Battle simulations and war games share common utilization characteristics and training advantages. They provide unit commanders, their staffs, and subordinate leaders the opportunity to practice critical command and control skills at greatly reduced cost and resource levels when compared to traditional field exercises.

These skills are developed and practiced in a simulated functional context in which players exercise their tactical roles during the course of a battle. They have to make decisions and exercise their command function in such critical areas as teamwork, tactics, maneuver, terrain utilization, and weapons systems employment and lethality, in response to constantly changing battlefield events. The wargames and battle simulations are intended to supplement field training at company and battalion levels and to constitute the principal operational training at those echelons where field exercises are not normally feasible.

With the escalating costs of conducting field exercises, it is likely that there will be increased reliance on battle simulations to train tactics and leadership skills. For this reason there has been a great deal of interest in determining their effectiveness for these purposes. Studies of the effectiveness of several of the current generation of battle simulations have been conducted to investigate what types of skills and principles are learned through experience with these simulations. One of these studies, which served

as the basis for the current research project, involved an experimental evaluation of the Dunn-Kempf board game battle simulation.¹

Dunn-Kempf is played at the company and platoon level, and is highly adaptable in terms of weapons systems employed, force composition, game rules, and tactical missions that can be performed. It provides leader training in such skills as tactical movement, direct and indirect fire engagements, use of smoke, battle planning, communication, and coordination of elements.

In the Miller and Bachta study, seven company level combined arms teams were trained on Dunn-Kempf using different training procedures. All teams used an active defense ARTEP mission and operated against a standardized enemy threat. The objectives of the research were to develop a means of measuring performance on the game, determine what people learn from the game, and identify the most effective method of play. The research resulted in the development of both subjective and objective performance measures which were used to evaluate training progress and alternative training strategies. The results showed that skill improvement was most striking in three skill domains: (a) relative priority assigned to high-threat targets; (b) coordination among team members; and (c) shifting of forces as the battle develops. It was concluded that an optimal method of playing the game, for training purposes, requires formal operations orders and communication restricted to tactically realistic channels, at least until the measures indicated that these two skills had been mastered.

The study described above and other effectiveness studies (CATTS,² CAMMS³) have evaluated performance and learning within battle simulation exercises. However, to date there has been no attempt to measure the effects of training in a particular battle simulation on performance in an operational training environment. The need now exists to determine if transfer of training can be measured in field tactical environments.

Measuring Transfer of Training

The study of transfer of training is concerned with the effects that learning skills in one type of setting have on performance in another setting.

¹ E.E. Miller and J.M. Bachta, "An Experimental Evaluation of a Tactical Game for Company Level Training," U.S. Army Research Institute, Draft Technical Report, January 1978.

² H.F. Barber and I.T. Kaplan, "Battalion Command Group Performance in Simulated Combat," U.S. Army Research Institute Technical Paper 353, March 1979.

³ I.T. Kaplan and H.F. Barber, "Evaluation of a Computer-Assisted Battle Simulation: CAMMS Versus a CPX," U.S. Army Research Institute Technical Paper 355, April 1979.

Functionally, transfer occurs over a continuum from positive to negative effects depending on such factors as the similarity of stimuli or cues within the two environments and the appropriateness of the learned response to the requirements of the criterion environment. The types of skills subject to transfer can range from simple psychomotor responses to the application of complex principles. Transfer of training is the ultimate test of training effectiveness where training is conducted outside of an operational setting.

To measure the transfer of training effects from a battle simulation to a field tactical environment three questions must first be answered. Namely: To what tactical environment might the learning transfer? What performance can be taken as evidence of transfer? How can the performance be measured? The questions of what performance to measure and how to measure it were addressed for the Dunn-Kempf battle simulation during the previous study. It still remained, however, to determine whether the measures and techniques which had been used in Dunn-Kempf could be used or adapted to a field tactical environment. This determination became the major focus of the present study.

The initial phase of this project analyzed candidate criterion environments (CCE) for measuring transfer of training from Dunn-Kempf.⁴ The analysis identified six types of field exercises available in USAREUR which could accommodate company level operations. The six identified were: the command post exercise (CPX); tactical exercise without troops (TEWT); field training exercise (FTX); Army Training and Evaluation Program (ARTEP) exercise; REALTRAIN; and REFORGER.

Transfer of training research literature was reviewed to identify important theoretical considerations for the selection of a criterion environment. From this review it was determined that factors of task commonality and functional similarity between environments would be necessary for transfer to occur, and that ability to measure transfer would be affected by the degree of control attainable over experimental conditions. Therefore three separate approaches were taken to the analysis of candidate criterion environments. The first analysis compared operational characteristics of Dunn-Kempf and the CCE such as ability of each to represent various combat assets, combat threats, combat effects, other battlefield characteristics, and the comparability of exercise control. The analysis of task commonality judged the potential applicability, in each of the CCE, of skills previously measured in Dunn-Kempf. The third analysis considered aspects of exercise complexity and control within the CCE as they could affect measurement of transfer.

REALTRAIN field exercises ranked highest on the analyses of functional similarity and task/skill applicability, but below the comparably sized field exercises (FTX, ARTEP) with respect to ease of control. Because the latter

⁴ G. Kress, et al., "Analysis of Tactical Training Environments for Transfer of Training Applications," Interim Report, Contract # MDA 903-78-C-2042, Human Resources Research Organization, January 1979.

consideration was judged subordinate to factors favoring transfer, REALTRAIN was selected for a criterion environment.

The overall results of the initial analyses described above suggested that the Dunn-Kempf measurement methodology needed to be further developed and shown to be usable in a field environment before transfer of training could actually be evaluated. Specific recommendations resulting from the analyses were: to utilize REALTRAIN as a criterion environment; to identify and specify parameters of exercises that could be conducted both in REALTRAIN and in Dunn-Kempf; to revise, develop, and refine measurement methodology which would apply to both types of exercises; and to try out the measures in the field to establish their utility and reliability.

PURPOSE

The purpose of the research reported here was to develop a methodology to measure transfer of training from a battle simulation (Dunn-Kempf) to a tactical field training exercise (REALTRAIN). This report documents the research activities leading to the development of this methodology. Specific research objectives were:

- (1) To determine empirically which measures previously developed for Dunn-Kempf evaluation can be applied to various missions in REALTRAIN field exercises.
- (2) To further develop and define performance criteria and measurement scales for tactical field performance evaluation.
- (3) To assess the reliability of measures when applied to field exercises.
- (4) To identify operational characteristics (e.g., mission, combat support assets) that must be controlled for both the Dunn-Kempf board game and REALTRAIN field exercises for a transfer of training experiment.
- (5) To design an experiment to measure transfer of training between Dunn-Kempf and REALTRAIN exercises.

METHOD

The sequence of tasks that were undertaken in developing the transfer of training methodology is depicted in Figure 1. Each of the seven steps is discussed in detail below.

1. Develop Approach. The components of a transfer of training methodology must include: a training system; a performance evaluation instrument; a means of applying the performance instrument; a criterion environment in which to measure performance; and an experimental design which prescribes a schema for conducting training and measuring performance so that the effects of training can be interpreted. For the present methodological development, REALTRAIN

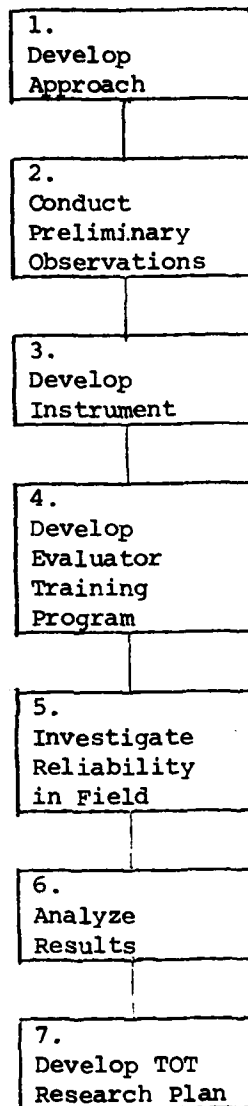


Figure 1. Research Activities for the Development of a Transfer of Training (TOT) Measurement Methodology

field exercises were to serve as a criterion environment, and the Dunn-Kempf measurement methodology was to be adapted for use in a field tactical environment as the means to measure transfer. Measuring transfer of training would require the evaluation and comparison, in REALTRAIN exercises, of experimental units that have participated in Dunn-Kempf training and control units that have not had Dunn-Kempf training. The approach to developing the methodology then was to identify and resolve problems faced in applying the Dunn-Kempf measures in the field.

The measurement methodology employed in the Miller and Bachta study used military evaluators to rate player performance in Dunn-Kempf exercises on forty measures of tactical skills. The measures were appropriate to an active defense mission which was the only mission used for evaluation. The ratings were made on a five point scale. Two problems were initially identified in applying this methodology for field evaluations. The first problem was that the company defensive mission requires a battalion sized opposing force (OPFOR). Because of its scale, realistic support of this mission in a field exercise was not feasible. Therefore the decision was made to base the transfer of training methodology on an offensive mission which would require only a platoon OPFOR.

The second problem recognized was that of measurement reliability. The reliance on trained military judgment to rate performance is necessitated by the fact that field tactical performance comprises highly complex behaviors occurring under highly variable conditions. These behaviors are not sufficiently definable or quantifiable to enable objective measurement. However, using subjective performance ratings creates a problem of unreliability. That is, differences in performance ratings made by different evaluators may reflect difference between raters rather than actual differences in the performance observed. The importance of measurement reliability is emphasized in the following statement from Jeantheau and Andersen (1966).⁵

"Reliability refers to the stability or repeatability of measurement. A proficiency test based on unreliable measures will not discriminate between individuals of different skill levels. In the evaluation process, the ability of a measure to differentiate between different levels of performance is critical."

The problem of reliability signaled a need to take every step possible to build stability into the measures used, and to test the finalized instrument for reliability before using it as the basis for a transfer of training experiment.

Therefore, approaches to address the problems of measurement reliability were developed. Reliability of measurement is normally achieved through

⁵ G.G. Jeantheau and B.G. Andersen, "Training System Use and Effectiveness Evaluation," Technical Report NAVTRADEVCEEN 1743-1, 1966.

standardization of measurement procedures. The approaches identified for the standardization of the instrument were: (a) to identify and define the components of tactical skills involved in each measure so that ratings would be made on the same basis; and (b) to explicitly define the anchor points on the rating scale used. The problem of standardizing application of the measures was addressed by developing an evaluator training program to explain the use of the instrument.

The approach to testing the reliability of the performance instrument requires that two evaluators be assigned to each player position over a series of exercises so that paired ratings could be analyzed statistically to derive a reliability index.

2. Preliminary Observations. Prior to adaptation of Dunn-Kempf measurement methodology for field data collection, a series of four REALTRAIN exercises were observed at the Baumholder training area. Observation of these exercises served several purposes:

- a. To learn how the REALTRAIN exercises are conducted and controlled in USAREUR, and to identify operational characteristics of the exercises for comparison to Dunn-Kempf.
- b. To monitor those positions involved in the play of Dunn-Kempf to determine if relevant performance could be observed.
- c. To identify and anticipate data collection problems and to develop data collection strategies to be considered during instrument development.
- d. To determine data collection personnel and material support requirements.

During the different exercises, project personnel assumed possible data collector positions, monitored the tactical net (radio channel), control net and the indirect fire control net. Two of the exercises were observed from a fire tower which afforded the best available view of the training area. Additionally, informal discussions were held with the leaders of the control cadre about options and levels of control over the exercises.

Information obtained relating to data collection procedures and problems is summarized as follows:

- a. It is normally feasible for data collectors to ride in an armored personnel carrier (APC) with the player they are evaluating or, as an option, to follow the vehicle in a jeep and monitor the appropriate tactical net. Traveling within the player vehicle enables closer observation of the player and access to non-radio communications, but limits capability to view surrounding activities. Traveling independent of the player vehicle can allow a fuller view of the coordination of tactical elements but limits direct

observation of the player. There are, however, severe limitations on space and visibility for an evaluator in a tank. Therefore, if the player is in a tank, following in a jeep is preferable.

b. Compared to Dunn-Kempf, mobility and maneuverability in the field can be relatively limited. Vehicles become mired in mud and there are normally restrictions limiting damage to timber and maneuver areas in general. In heavily wooded areas, tracked vehicles must generally use existing trails.

c. Observation of tactical exercises can be restricted for a number of reasons. The type and nature of terrain is a significant factor. Weather conditions, such as fog, can limit visibility. Thus members of the element may not always be in visual contact. Such factors of visibility can affect both performance of players and the ability to evaluate performance. Because of limitations on observation by the data collectors, access to radio traffic and knowledge of offensive plans are very important to interpretation and evaluation of tactical behaviors.

d. Logistical problems may interfere. Vehicles break down. Communications are subject to interruption. These are problems not encountered by players or evaluators in Dunn-Kempf exercises.

e. Another problem was discovered which is not encountered in Dunn-Kempf. In Dunn-Kempf, vehicles become casualties, but the players continue to control the remaining vehicles. In REALTRAIN, the commander, platoon leader, or FIST chief can become a casualty and preclude further evaluation of the affected player.

Information was also obtained about the operational characteristics and control of REALTRAIN exercises.

a. REALTRAIN can accommodate any unit configuration and provide for casualty assessment for all organic weapon systems. Indirect fire support, in terms of mortar, artillery, and smoke missions, is played effectively. Close air support and air defense artillery are not normally playable.

b. Direct fire suppression is playable in REALTRAIN but players tend not to use it because their own weapon's signature may disclose their position. Functionally, most engagements are one-on-one.

c. Scenario control in an offensive mission is limited to (1) placing restrictions on terrain use by marking boundaries or barriers, (2) specifying locations of the assembly area and the objective, and (3) determining general guidance in the Operations Order (OPORD). Because the offense is in an active role while the defense is in a reactive role, it is more difficult to "standardize" an offensive mission. However, there are more options open to leaders during an offense, and more activity and behavior to observe.

In addition to field observations, a meeting was arranged with company personnel who were familiar with both Dunn-Kempf and REALTRAIN to discuss the relationship between the two types of training. Topics discussed included: How to collect certain types of information in REALTRAIN, and the advisability of using REALTRAIN controllers as data collectors; the differences between offense and defense as they affect certain tactical behaviors such as use and nature of communications, fire distribution, and use of indirect fire; and, the type of tactical principles that they felt were enhanced by the use of Dunn-Kempf. Much of the information generated in these discussions and observations were used to guide the development of the performance evaluation instrument and the planning for data collection.

3. Develop Instrument. Development of the Performance Evaluation Instrument proceeded through three stages. These stages were: (a) design of a prototype instrument for an offensive mission; (b) interactive field tryouts and revisions; and (c) a final refinement stage.

a. Instrument design. The first step involved reviewing the original Dunn-Kempf measures (Appendix B, Table B-1) to identify those measures which could be used or adapted for an offensive tactical exercise conducted in a REALTRAIN field environment. The two main criteria used to screen each item were:

(1) The measure must have a high transfer of training potential. Ten of the original forty Dunn-Kempf measures were shown to register little change over four exercises and were designated in that study as low-gain items. (In some cases this was due to a ceiling effect in which performance was high to begin with and there was little room for improvement.) It was reasoned that the probability of measuring transfer in a criterion environment was low for skills which showed little gain in Dunn-Kempf. Therefore, the low-gain items were eliminated. A second type of measure was eliminated using the same criterion. Several of the Dunn-Kempf measures pertained to crew served weapons (primarily HAWs). In Dunn-Kempf these measures evaluated platoon leader decisions or actions, but in REALTRAIN they would apply to the weapon crews themselves, independent of the platoon leader input. An example is, "Did HAWs coordinate fire with tank direct fire suppression so as to minimize vulnerability of HAW to enemy counter fire?" In Dunn-Kempf, the platoon leader would coordinate these weapon systems but in REALTRAIN, the HAW crew would be responsible for this activity. Therefore, those items were deleted or revised to evaluate guidance/control from the platoon leader level.

(2) The measure must apply to an offensive mission. Nineteen of the thirty high-gain and moderate-gain Dunn-Kempf measures applied essentially to a defensive operation. An example is, "Were registered fires located along most likely enemy avenues of approach?" Some of the nineteen measures however embodied tactical principles which were also applicable to the offensive mission although the measure itself was not. As an example, "Did platoon positions selected by the CO afford/provide concealment?" The principle and skills of using cover and concealment apply to any type of mission. Therefore, in such cases, the general tactical princi-

ples in the measures were identified and retained for incorporation in measures for the offense.

The result of the screening was that only ten of the Dunn-Kempf measures appeared usable in nearly original form.* Therefore, it was necessary to identify gaps in the coverage of the performance involved in an offensive mission. Offensive task listings were generated from the three offensive missions (movement-to-contact, hasty attack and deliberate attack) contained in the company level combined arms ARTEP 71-2 (Chapter 8, Appendices 10-12). Tasks specified in the accompanying Training/Evaluation Standards. From this listing and from tactical principles extracted from deleted Dunn-Kempf measures, skill areas were identified which generalized across offensive missions. The additional skill areas identified were:

- . maneuver planning
- . fire planning to support maneuver
- . movement techniques
- . use of FRAG orders
- . reaction to contact
- . use of fire and maneuver after contact
- . control of fire distribution
- . overall use of indirect fire (IF)

The next step was the development of measures for a prototype performance instrument which tapped the skill areas identified. Construction of an instrument to evaluate tactical performance is a judgmental process, particularly with regard to the level of detail at which to focus. To guide this process, three basic principles were followed:

- . Measures should be sufficiently general to apply to all offensive missions and scenarios.
- . The instrument should be sufficiently broad in scope to cover all leader functions (pertaining to tactics) at each position evaluated.
- . The measures should deal with all likely events (and possible reactions to events) in an offensive scenario that pertain to the tactical skill area measured.

In construction of individual measures from the tactical skill areas identified, the requirement for item stability was a major consideration. An approach commonly used for individual procedural tasks is to define behavioral criteria for each measure in terms of the specific procedures being measured. After some investigation, this approach was deemed impractical for application to tactical skills. Tactical behaviors occur under a wide range of circumstances, and there are many possible variations in their execution relative to the existing circumstances. For these same reasons, doctrine lacks specificity in defining tactical behaviors, and offers only representative examples. Therefore it was not possible to

* Appendix B, Table B-1 measures 1, 3, 6, 8, 10, 13, 19, 25, 27, and the OPORD checklist.)

define criteria in specific behavioral terms for each measure that would apply every time the appropriate skill was required.

As a resolution to this problem, generalized performance criteria were developed for the rating scale which accompanies each measure. Definitions of satisfactory and unsatisfactory performance were developed and anchored to two points on the five point rating scale. The other three points were defined in terms of these two points. The rating scale used and definitions of points are shown in Exhibit 1. The rating scale, as shown, provides a general definition for "very poor," "unsatisfactory," "borderline," "satisfactory," and "very good" performance for each of the skills being evaluated. The evaluator has to apply these criteria in light of the circumstances which prevail at the time the behavior is being observed.

Application of the general performance criteria requires judgments about the appropriateness of player actions and their correspondence to doctrine. To provide a basis for relating the generalized criteria to specific skill areas, the evaluator is required to answer questions about components of skills (identified from doctrine) and situational factors relevant to the particular skill. For example when evaluating the fire plan, the evaluator is asked to determine whether smoke missions are planned to mask movement through open areas, or if preplanned fires are designated on suspected enemy positions. These questions identify both doctrinal skill components (use of smoke, preplanned fires) and situational factors (open areas, enemy positions). Depending on the measure, subelements may consist of a listing of relevant procedural steps or an itemizing of significant factors which should be considered in executing and in evaluating a tactical skill. Where behaviors are of a repetitive nature, such as movement, execution of overwatch, and communication, the data collection forms were designed to require notations on the pertinent facts each time the behavior was observed. Observations and judgments about these subelements of skill domains are generally easier to make than evaluations of a global measure, and at the end of the exercise these notations represent a summary of all performance observed to aid the evaluator in assigning an overall rating. The rationale behind this approach was that stability of a measure would be enhanced by requiring all evaluators of a player position to consider the same elemental factors when determining a rating.

For the period of instrument development, a further dimension was added to the use of the rating scale. A ten-point confidence scale (ranging from 0-100%) was positioned below each rating scale so that the evaluator could indicate his confidence in the accuracy of the rating based on his ability to observe performance and his understanding of the measure. The purpose of the confidence index was to help in establishing the utility of the measures during a planned period of field tryouts. The intent of the index was to draw attention to items which evaluators had difficulty in rating. It was reasoned that if there were measures that evaluators were not confident in rating, these would probably detract from overall instrument reliability. The confidence scale and instructions are presented in Exhibit 2.

EXHIBIT 1: Tactical Performance Rating Scale

very poor	unsatisfactory	borderline	satisfactory	very good

The points on the scale are generally defined as follows: (Note that not all of the points in a definition will always apply to a measure.)

"BORDERLINE" - This is the balance point on the scale. It means that the overall performance on the measure could not consistently be judged satisfactory or unsatisfactory. The positive and negative behaviors cancelled each other out.

"SATISFACTORY" - Means that most of the time the behavior that was appropriate was performed, the performance met accepted standards of tactical doctrine; it was in compliance with orders/directives; it was appropriate to the mission, terrain, and enemy situation.

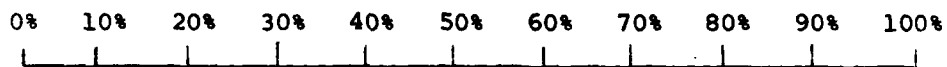
"VERY GOOD" - Means that almost all of the time, the requirements for a "satisfactory" rating were met, and in addition, the timing and execution of performance were clearly superior.

"UNSATISFACTORY" - Means that most of the time the behavior that was appropriate was not performed; or if performed, it did not meet accepted standards of tactical doctrine; or performance was not in compliance with orders/directives; or performance was not appropriate to the mission, terrain, and enemy situation.

"VERY POOR" - Means that almost all of the time the tactical performance was unsatisfactory in terms of the points mentioned above and/or showed some glaring deficiencies.

EXHIBIT 2. Instructions for Use of Confidence Index

CONFIDENCE INDEX - For each performance rating given you will also indicate how confident you are in the accuracy of the rating. This confidence rating should be based only on how much you were able to observe and how well you were able to interpret the measure as it applied to what you observed. The confidence rating can vary from 0% to 100% according to the following scale:



An example of a complete item is shown in Exhibit 3. The skill domain of interest in this measure is the platoon leader's ability to minimize vulnerability of his element to enemy detection and fire during movement. The global measure is presented in the form of a question: "Did the bounding element minimize exposure during movement?" The skill components identified for consideration are: Selection of the movement technique which is appropriate to the situation, the use of artificial or natural cover to mask movement, the use of speed where cover is not available, the use of vehicle dispersion at all times, and ending movement in a covered and/or concealed position. These separate skill components are to be judged each time the platoon leader's element executes a bound (up to ten bounds). Then based on the record of the series of observations, the evaluator will make an overall rating of the movement technique of the platoon leader's element using the rating scale and general performance criteria. If, for any reason, the evaluator is not able to observe enough to make these judgments, he can indicate this fact and explain the reason. This option is routinely included in all items and is used as a way to document problems encountered in field observation.

To the extent possible, the measurement items were organized chronologically to proceed through the general exercise phases of planning, movement, contact, fire and maneuver, and assault. The prototype performance evaluation instrument which was developed contained the measures presented in Table 1. This instrument was then subjected to a period of thorough field trial and revision, described below.

b. Field tryouts. The second major state of instrument development consisted of field trials and revision of the prototype instrument. An ad hoc committee of five officers and one NCO was assigned to this project for a five-day period. There were several objectives for the field trials:

- . to determine the applicability of the measures to the field exercise;
- . to try out and optimize data collection procedures;
- . to try out and refine data collection forms;
- . to obtain a critique of the instrument by military personnel with respect to the importance of the tactical behaviors selected for measure; whether there were significant aspects of tactics not being measured; and finally, whether the approach or focus of measures was appropriate.

The first day of this period was devoted to: (a) a terrain reconnaissance and an orientation to the objectives of the project; (b) the particular purposes of the tryout period; and (c) the roles/functions to be served by the committee members. Then the group reviewed the instrument in detail to achieve familiarity with the behaviors to be observed, the events to be recorded, and the use of recording formats.

EXHIBIT 3. Example Performance Measure

PL

6. Did the bounding element minimize exposure during movement?

Bound #	Movement technique utilized: (T,TO,BO)	Was technique appropriate to situation? (Y,N)	Was smoke used to mask bound through open areas? (Y,N, <u>Not Available</u>)	Was available terrain and foliage used to mask movement? (Usually, Sometimes, Rarely)	Did the element use rapid movement across open areas? (Y,N)	Did they maintain vehicle dispersion? (Y,N)	Did they bound to covered/concealed positions? (Y,N, <u>Not Available</u>)
1							
2							
3							
4							
5							
6							
7							
8							
9							
10							

a) RATING: (Element rated _____) Did the bounding element minimize exposure during movement?

very poor unsatisfactory borderline satisfactory very good

Confidence level in accuracy of evaluation:

0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

b) It was not possible to observe enough to make an evaluation: _____

Reason: _____

TABLE 1

Dunn-Kempf/REALTRAIN Measures

(Prototype)

	<u>Position</u>
* 1. Operations Order	CO
2. Was the scheme of maneuver appropriate to the mission and terrain?	CO
3. Was Fire Plan developed which would support the scheme of maneuver?	CO/FIST
4. Was the LD crossed on time?	CO/PL
5. Was the movement covered by overwatch?	CO/PL
6. Did the bounding element minimize exposure during movement?	CO/PL
7. During movement phase, did individual vehicles bound to covered and/or defilade positions?	ALL
8. Were overwatch positions well selected?	CO/PL
9. Were vehicles in overwatch elements in covered and/or defilade positions?	CO/PL
* 10. Was CO kept informed of friendly situation?	CO/PL
* 11. Was CO kept informed of enemy situation?	CO/PL
12. Was the FRAG Order complete and appropriate to the situation?	CO
13. When a bounding elements was fired upon did the overwatch element take appropriate/prescribed action?	CO/PL
14. When the bounding elements was fired upon did the overwatch element take appropriate/prescribed action?	CO/PL
* 15. Did TOWs give priority of fires to most dangerous enemy weapons?	PL
* 16. Tactical Concept: Shoot and Move.	ALL
* 17. Did TOWs coordinate fire with tank direct fire suppression so as to minimize vulnerability to enemy counter fire?	PL
18. Are units coordinated and brought into position for assault/hasty attack using appropriate fire and maneuver techniques?	CO/PL
19. Was suppressive fire used to support the final assault?	CO/PL

- | | | |
|-------|---|---------|
| 20. | When contact was made, did platoon leader control the fire distribution of his elements? | PL |
| * 21. | Did platoon leaders know the location of registered and preplanned fires? | PL |
| * 22. | Was ARTY used to suppress the most dangerous enemy weapons? | FIST |
| * 23. | Were ARTY adjustments quick and accurate? | FIST |
| 24. | Was maximum use made of registration points (RP)/preplanned fires (PF)? | FIST |
| * 25. | Were registration points (RP) and preplanned fires (PF) located so as to be easily adjusted from? | CO/FIST |
| 26. | Was maximum use made of indirect fire support? | FIST |

* Items retained from original Dunn-Kempf measures (Appendix B)

The instrument was then tried out on five company-level REALTRAIN exercises over the next four days. After each day in the field, the committee reconvened in garrison to discuss data collection problems and potential solutions, and to consider recommendations for revisions to the instrument. Types of revisions that were incorporated through this process included: Changing the echelon at which an observation was made (i.e., Platoon Leader to CO, or the reverse); changing the informational elements making up a measure; revising the recording or coding of observations; and excluding measure as not relevant or collectable in REALTRAIN.

As a result of the tryout, the 26 measures were reduced to 17, either through elimination or consolidation. Referring to Table 1, items 15, 17, and 23 were eliminated as unobservable in REALTRAIN. Measure 23, "Were ARTY adjustments quick and accurate?" provides an example. Both accuracy and time to adjust are partially dependent on control personnel in REALTRAIN and this effect could not be isolated from the player's (FIST chief) performance. Six pairs of items were consolidated into single measures. The consolidated items related to the same skill domains, and consolidation served to simplify observation and recording procedures for the evaluators. For example, the two measures dealing with reports on friendly and enemy situations were consolidated into a single collection form from which the two separate ratings could be made. Items consolidated in this manner were 6 and 7, 8 and 9, 10 and 11, 13 and 14, 19 and 22, and 24 and 25.

For the final exercise of the field trials, two data collectors evaluated each player position. This technique was tried out for later use in an investigation of instrument reliability. The pairs of evaluators for the FIST and platoon leader positions rode in APCs and the two CO evaluators observed the exercise from the same jeep. This arrangement was found to cause no data collection problems nor interference with the exercise. However, it was necessary to instruct the collocated evaluators to arrive at their ratings independently. The data collector positioning described above also turned out to be the most effective for one-on-one evaluation. The CO evaluator obtained a broader perspective by moving independently from the CO's track and monitoring the tactical net. Traveling in the player vehicle was deemed more appropriate at the platoon leader level except for an armor platoon where following by jeep was almost mandatory. Because of the information required for FIST evaluation, it is preferable that the evaluator be located in the FIST track.

While performance data collected during the period of field trials was incidental to the primary purpose of instrument development, a summary of the findings with respect to the confidence index is of interest:

- . Averaged by position, the highest confidence index was for the FIST evaluator, the lowest for the CO evaluator.
- . The confidence index ranged from 73-91 (average = 82.5) across data collectors and exercises for retained items. The range for deleted items was 20-90 (average = 71.6).
- . The average confidence index did not appear to increase across exercises. This was probably due to the fact that the instrument was under revision, and therefore changed on successive trials.

c. Instrument refinement. The third and last stage of instrument development entailed: Review of recorded discussions with the ad hoc committee; review of tactical doctrine to verify the consistency of revisions determined at Baumholder; and, incorporation of the changes into a new document. The format for some items was simplified during this period, and the instructions accompanying the instrument were elaborated.

The player position in the prototype instrument for which there were the fewest measures was that of the FIST chief. Some of the original Dunn-Kempf measures dealt with the use of direct fire, but the FIST team concept was not in use at the time of the previous study. As a result of the field tryout of the instrument the number of measures of FIST performance was reduced from six to four (compared to 12 and 14 for the CO and PL respectively). Therefore, an effort was made to expand this area of evaluation.

An artillery representative with the Force Modernization Division of DCSOPS, HQ, USAREUR was enlisted to review the instrument and suggest additional aspects of FIST operation to evaluate. Through this source, recently available publications were obtained on artillery doctrine along with a (yet unpublished) task analysis of the FIST position developed at Fort Sill. Through discussions and a review of this material, existing FIST items were expanded and five other measures were modified to encompass FIST responsibilities.

An item by item explanation of the evolution from the original Dunn-Kempf measures to the finalized performance evaluation instrument is presented in Appendix C. The performance evaluation instrument that was developed is presented in Appendix A. In comparing the final field evaluation instrument with the instrument used in the Dunn-Kempf study, the following differences should be noted:

- . The current instrument is designed to evaluate an offensive rather than defensive mission.
- . The instrument now identifies the player position which is the object of each measure (In fact, separate data collection packets are composed for each position).
- . The instrument identifies components of tactical skills and pertinent situational factors involved in each measure.

- . The instrument identifies requirements for coordination between evaluators to share information and thus provide more complete perspectives for evaluation (this was militated for in adapting the Dunn-Kempf instrument to the large scale of a field evaluation).
- . The rating scale now anchors points to generalized criteria definitions of satisfactory and unsatisfactory performance.

4. Develop Evaluator Training Program. The performance evaluation instrument depends on the judgments of knowledgeable military personnel. While some commonality of background can be assumed, judgment is a very individual process. Therefore, an important aspect of the evaluation methodology developed was the training of evaluator personnel relative to their roles and the instrument they would be using. The orientation of the ad hoc committee members described in the previous section served as a model for the training program. The training program is intended for presentation the day prior to the first exercises for evaluation.

The approach to developing the program was to analyze what evaluators need to know in order to accomplish the purposes of the evaluation. First, evaluators should have a basic understanding of the research objectives and how those affect the focus of the evaluation. In addition, evaluators should understand basic REALTRAIN procedures, they need to be familiar with the terrain used for the exercises, they need to be familiar with the content of the measures (and the doctrine which pertains to them), the data required by the measures, and the recording procedures and forms. Therefore, the training program was planned around these requirements. (An outline/agenda of the program is included as Exhibit 4).

The training program begins with an explanation of the project and the evaluator's role in fulfilling project requirements. The remainder of the morning session covers the REALTRAIN concept and procedures, a map orientation, and a terrain walk. The afternoon session is planned as a discussion of the intent of the evaluation, an explanation of the instrument instructions, and an item-by-item review of the instrument. Discussion of the instructions includes an explanation of the use of the rating scale and definitions of the anchor points. It is emphasized that evaluations should be based on the circumstances that prevail at the time of observation. The rater is directed to take into account how well an individual performs on a measure in light of the assets and information he has, and the constraints and conditions that he faces. Therefore, the evaluator must keep constantly abreast of the developing situation.

The item-by-item review is handled in a group discussion format and includes an identification of the relevant tactical concepts, review of pertinent doctrine, and a discussion of the way situational variables can affect performance. Examples of the latter are elicited from the group. For each measure, the evaluators are asked to describe what they will look for, what they will base their judgments on, and what they would consider as good or poor performance. A group discussion of each item

EXHIBIT 4. Evaluator Training Program Outline

a.m. Orientation to Purposes of Evaluation

Evaluation of five leadership positions; focus on how each performs tactical skills.

Evaluator Role and Responsibility

Reliance on military judgment of tactical proficiency.

Terrain Orientation

Map reconnaissance; identification of assembly area, objective, and alternative routes of advance.

Terrain reconnaissance.

-- NOON RECESS --

p.m. Use of Performance Evaluation Instrument

General Instructions; read out loud and discuss.

Structure of Measures

Informational items; repeated observations.

Conditions for observation; coordination of information.

Rating Scale; point definitions and use of scale.

Situational Dependence; concept of application of principles; examples of how different circumstances affect performance; evaluation based on awareness of prevailing circumstances.

Review of measures; for each item consider:

- what behavior to observe
- when to observe
- how to observe (use of radio, positioning)
- what is tactical principle involved?
- what does doctrine say?
- what constitutes good performance?
- what constitutes UNSAT performance?

(SUGGESTION: Use of terrain board with miniature vehicles to set up and critique hypothetical situations)

Administrative Information

Vehicle/personnel assignments.

Call signs; company, platoon, and control frequencies.

is encouraged in order to achieve a consensus relative to levels of performance.

Finally, logistical and administrative information must be covered, such as vehicle assignments, call signs, control and tactical frequencies, and provisions for the evaluator group to confer after the exercises.

5. Investigate Reliability in Field. It was originally intended to test the reliability of the evaluation instrument in field exercises. However, available support, in terms of company-level REALTRAIN exercises and dedicated evaluator personnel, was very limited and was not adequate to meet the requirements of the reliability investigation. Therefore, this effort was restricted to a field tryout of the revised instrument. REALTRAIN exercises conducted at Hohenfels were used to support the field tryout. Dedicated evaluator personnel were not available; however, the REALTRAIN controllers assigned to certain of the Dunn-Kempf player positions undertook the performance evaluation as a secondary function. Reliability of measures could not be studied during these exercises because it was not feasible to shift evaluators from control assignments in order to get a second evaluation of each player position. A further limitation existed in that FIST personnel acted in a player/controller capacity for these exercises and were not accessible to the evaluation group (FIST personnel rotated with the companies that were involved in the exercises, were not available for orientation, and in essence would have had to evaluate themselves). Moreover, CO-level evaluators were coordinating the operation of all vehicle controllers under their supervision, monitored only the control net, and traveled independently of their player counterpart, so data at this level was very sparse. Therefore, this was effectively a tryout of the platoon leader portion of the instrument. Under these limited conditions the instrument was applied to six exercises conducted during this period.

6. Analyze results. The field tryout of the revised instrument was limited to the items rated at the platoon leader level which involved 14 of the 18 instrument measures. Primary interest during the tryout was centered on the frequency with which items could be rated, the confidence the evaluator had in ratings, and the reasons cited for not rating items. The results obtained consisted of:

- a. Frequencies with which measures were rated, for all exercises, expressed as a percentage of potential opportunities for rating. The potential opportunities were calculated as the number of exercises times the number of evaluators for each exercise.
- b. Tabulation, for each measure, of the number of exercises during which at least one rating was obtained.
- c. Tabulated ratings and confidence indexes (CI) for each measure across exercises, and an average CI for each measure based on all exercises.

A summary of findings is shown in Table 2. The average percentage with which measures were rated was 67% of the theoretical potential for observation. All but four measures were rated at least once per exercise and all but one were rated in at least five of six exercises. Six items (measures #8, 10, 12, 13, 14, and 17) were identified as having a low collection frequency and/or low average confidence index (below 60 percent on either scale).

The reasons for not rating items fell into four general categories. The most common reason given (12 instances) is that the platoon leader became a casualty before a number of the relevant skills could be exercised and evaluated. (Because of the potential impact on evaluation in a transfer of training experiment, it may therefore be necessary to "revive" a player if he becomes a casualty early in the exercise). A second reason (six instances) was that the density of woodland in the particular training area limited observation of the tactical behaviors being measured. The third category, which could also be terrain related, is that the relevant technique (in all six cases, "overwatch") was not employed. The preceding two reasons point to the need, when selecting terrain sectors for exercises, to consider whether the various tactical techniques can be applied. The final category of reasons pertained to cases where the rated player's element had not made contact with the enemy before the exercise was halted (four instances). This is an understandable circumstance for the rear platoon in an offensive operation.

The overall frequency of rating (.67) seemed somewhat low. Therefore, it was compared to data from exercises at Baumholder where a dedicated team of evaluators had been used. It was found that the average frequency of obtaining ratings over five exercises there was 0.65. These frequencies appear quite consistent. Therefore, given the difficulties inherent to field evaluations, it is probably not reasonable to expect a much higher rating frequency.

The average confidence index (CI) at Hohenfels was 64 (on a scale from 0 - 100). This compares to a CI of 82.5 for exercises at Baumholder. Further comparison of rating forms between the two sets of exercises was undertaken to investigate this difference. Review of informational items making up various measures showed that, on the average, ratings of tactical skills were made on the basis of fewer observations at Hohenfels. This is not unexpected in view of the shared control responsibilities of evaluators there, which may have also contributed to the lower confidence indexes. These findings militate for the use of a dedicated evaluator team to perform the evaluations. This interpretation is further supported by review of measures with low rating frequencies or low confidence indexes.

The six measures with lower than 60% rating frequency or confidence indexes were singled out as presenting potential measurement problems. These measures cover such tactical domains as reaction to enemy contact, use of fire and maneuver, use of suppressive fire and indirect fire, etc. All six measures are evaluated in the context of enemy engagement. This

TABLE 2

FREQUENCY OF RATINGS AND AVERAGE CONFIDENCE INDEXES
FOR MEASURES USED IN HOHENFELS EXERCISES

<u>Measure #*</u>	<u>a. Ratings/ Opportunities</u>	<u>b. # of Exercises During Which Measure Was Rated (Out Of 6)</u>	<u>c. Average Confidence Index</u>
3'	14 / 15 = .93	6	(N/A)
5	13 / 15 = .87	6	69
6	13 / 15 = .87	6	65
7	30 / 45 = .67	6	71.5
8'	23 / 30 = .77	6	54
10	16 / 30 = .53	5	57
11	9 / 15 = .60	6	67
12	8 / 15 = .53	6	66
13	5 / 15 = .33	3	63
14	7 / 15 = .47	5	64
15	15 / 15 = 1.00	6	(N/A)
16	9 / 15 = .60	6	64
17	7 / 15 = .47	5	64
18	12 / 15 = .80	6	62
All items	181 / 270 = .67		64

*Items are referenced to the Performance Evaluation Instrument at Appendix A.

is the same period when REALTRAIN control functions are most demanding. A reasonable interpretation then was that REALTRAIN controller duties interfered with observation and evaluation of these particular items. To further investigate this interpretation, the frequency and confidence with which the comparable items were rated at Baumholder was reviewed. Four of the six items had rating frequencies over 75% and confidence indexes over 80. For these items at least the interpretation appeared to be supported.

For two measures, #13 and #17, the problems of low frequency were also found at Baumholder (38% and 50% respectively). Both of these items had been revised since the Baumholder exercises to utilize input from the FIST evaluator, and could not be fully evaluated at Hohenfels. Determination with respect to the utility of these measure will require additional monitoring and possible adjusting during any further use of the instrument.

The remainder of the measures appeared to present no particular problems. Understandably the two items with the highest frequency (over 0.90) are those collected before the exercise begins. The fact that most measures were rated during all exercises is a favorable finding for use of the instrument in a transfer of training experiment. At the platoon leader level, there is a high likelihood of obtaining at least one rating per exercise because there are three platoon leaders and three platoon leader evaluators. This is, of course, not the case for the positions of CO and FIST chief. However, most of the measures (12/18) are evaluated at more than one position, so the likelihood of obtaining ratings for most measures for each exercise during a transfer of training experiment appears to be good.

7. Develop a Transfer of Training Research Plan. A research plan for a transfer of training experiment has been prepared and is attached to this report. The research plan outlines the experimental design, specifications for the conduct of Dunn-Kempf and REALTRAIN exercises, resource requirements and data analysis procedures. The plan identifies operational exercise characteristics which must be held constant for Dunn-Kempf and REALTRAIN exercises in terms of mission, OPFOR/friendly force ratios, unit configurations to include organic combat assets, available inorganic assets, and representation of external agencies (e.g., battalion CO/S-3, fire support coordinator (FSCoord), adjacent units). The actual design of the exercises would be effected in coordination with participating battalion training management personnel. Because the investigation of instrument reliability was not completed, it is included as the initial subtask of transfer of training experiment.

In addition to the Performance Evaluation Form, an Exercise Record Form is included with the research plan. The principal purpose of this form is to record casualty assessment data from the exercises. This

data can be used as a secondary performance index for comparison of experimental and control groups as it was for the preceding Dunn-Kempf study.⁶ Use of the form is explained in the research plan.

CONCLUSIONS

Based on the experiences from developing a performance instrument and two periods of field evaluations, the following conclusions have been reached:

1. Evaluation of tactical performance in a field environment is highly complex and relies heavily on well trained and experienced evaluators. The larger the scale of the exercise, the greater the amount of uncertainty that is accommodated in the evaluation process.
2. A dedicated evaluator team should be used for the study of transfer of training.
3. There is a need for consistent, thorough evaluator training preparatory to evaluations to help control for individual differences in approach to the subjective measures.
4. Probably not all measures will be rated in each exercise under the best of conditions.
5. There are many uncontrollable factors which may affect performance and evaluation of performance in a tactical field environment. The question of inter-rater reliability under these circumstances is a crucial problem facing all studies of the effectiveness of tactical training.

⁶

Miller and Bachta, op. cit., p. 25

T2/DCI-1

APPENDIX A

Dunn-Kempf / REALTRAIN

Tactical Performance Evaluation Form

Exercise # _____

Data Collector Name: _____

Date: _____

am _____ pm _____

Company Observed: _____

Element Observed: _____

Position(s) Observed: _____

GENERAL INSTRUCTIONS

THE PURPOSE OF THIS DATA PACKET IS FOR USE IN ASSESSING THE TACTICAL PERFORMANCE OF A COMPANY TEAM IN A MOVEMENT TO CONTACT/ATTACK SITUATION. THE PRIMARY FOCUS FOR EVALUATION IS ON THE UNIT LEADERS (CO, PLT. LEADERS, FIST/FOs) AND HOW THEY EXERCISE THEIR COMMAND AND CONTROL FUNCTIONS. WE ARE INTERESTED IN WHAT THE UNIT LEADERS DO, HOW WELL THEY DO IT, OR WHAT THEY FAIL TO DO. INDIVIDUAL INFANTRYMEN OR VEHICLE CREW MEMBERS ARE NOT BEING DIRECTLY EVALUATED. FOR EXAMPLE, IF A UNIT LEADER WERE TO ASSIGN TARGETS FOR HIS UNIT, WE ARE INTERESTED IN EVALUATING THIS CONTROL FUNCTION BY ITSELF RATHER THAN EVALUATING WHETHER INDIVIDUALS ACTUALLY HIT THE TARGETS THEY ARE SHOOTING AT.

EACH ONE OF THE ITEMS (TACTICAL PERFORMANCE MEASURES) IN YOUR DATA COLLECTION PACKAGE WILL USUALLY INVOLVE THE FOLLOWING:

- INFORMATIONAL QUESTIONS that relate to the performance being observed. These questions should be answered/checked during the course of the tactical exercise. The questions usually pertain to factors or aspects of tactical performance that are emphasized in Army doctrine.
- OBSERVATION CONDITIONS are brief statements of when to observe or how to observe the performance being evaluated.
- PERFORMANCE RATING SCALE - toward the end of each exercise you will rate the overall performance of the individual or section that you are observing on each of the performance measures contained in the data collector package. For all measures, information recorded during an exercise should be reviewed and considered when rating overall performance on each measure. Each of the ratings will be made using the five point scale which follows.

very					very
poor	unsatisfactory	borderline	satisfactory		good

The points on the scale are generally defined as follows: (Note that not all of the points in a definition will always apply to a measure.)

"BORDERLINE" - this is the balance point on the scale. It means that the overall performance on the measure could not consistently be judged satisfactory or unsatisfactory. The positive and negative behaviors cancelled each other out.

"SATISFACTORY" - means that most of the time the behavior that was appropriate was performed, the performance met accepted standards of tactical doctrine; it was in compliance with orders/directives; it was appropriate to the mission, terrain, and enemy situation.

"VERY GOOD" - means that almost all of the time, the requirements for a "satisfactory" rating are met, and in addition, the timing and execution of performance are clearly superior.

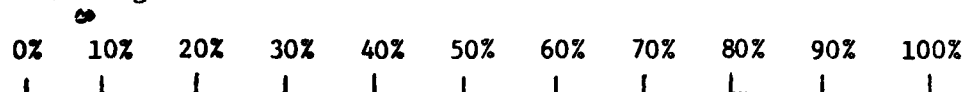
"UNSATISFACTORY" - means that most of the time the behavior that was appropriate was not performed; or if performed, it did not meet accepted standards of tactical doctrine; or performance was not in compliance with orders/directives; or performance was not appropriate to the mission, terrain, and enemy situation.

"VERY POOR" - means that almost all of the time the tactical performance was unsatisfactory in terms of the points mentioned above and/or showed some glaring deficiencies.

• CONFIDENCE INDEX - for each performance rating given you will also indicate how confident you are in the accuracy of the rating. This confidence rating should be based only on how much you were able to

observe and how well you were able to interpret the measure as it applied to what you observed.

The confidence rating can vary from 0% to 100% according to the following scale.



AT THE END OF MOST MEASUREMENT ITEMS THERE IS AN OPTION WHICH ALLOWS YOU TO INDICATE IF IT WAS NOT POSSIBLE TO OBSERVE ENOUGH TO RATE PERFORMANCE ON A MEASURE. IF THAT IS THE CASE, WE WOULD LIKE YOU TO IDENTIFY THE REASON THAT OBSERVATION OR MEASUREMENT WAS NOT POSSIBLE DURING THAT EXERCISE.

IN ALL INSTANCES WE WOULD ALSO LIKE YOU TO WRITE DOWN REMARKS ABOUT ANY CIRCUMSTANCES, WHICH YOU CONSIDER UNUSUAL OR EXCEPTIONAL, THAT AFFECTED PERFORMANCE.

Dunn-Kempf/REALTRAIN Measures

	<u>Position</u>
1. Operations Order	CO
2. Was the scheme of maneuver appropriate to the mission and terrain?	CO
3. Was fire plan developed which would support the scheme of maneuver?	CO/FIST
3.' Were direct fire weapon systems assigned priority of targets?	PL
4. Was the LD crossed on time?	CO
5. Was movement covered by overwatch, when appropriate?	CO/PL
6. Did the bounding element minimize exposure during movement?	PL
7a. Did overwatch positions selected maximize observation ability?	PL/FIST
7b. Did the overwatch positions selected enable fast and effective fire support for the maneuver elements?	PL/FIST
7c. Did the overwatch positions selected minimize vulnerability to enemy detection and fire?	PL/FIST
8a. Was CO kept informed of friendly situation?	CO
8b. Was CO kept informed of enemy situation?	CO
8'a. Was CO kept informed of friendly situation?	PL/FIST
8'b. Was CO kept informed of enemy situation?	PL/FIST
9. Was the FRAG Order complete and appropriate to the situation?	CO
10a. When engaged, did the engaged element take the appropriate action?	PL
10b. When engaged, did the supporting element take the appropriate action?	PL
11. Did weapon systems move after engaging enemy targets?	PL
12. After enemy contact is made, are elements controlled using appropriate fire and maneuver techniques?	CO/PL
13. Was suppressive fire used to support the final assault?	CO/PL

	<u>Position</u>
13.' Was suppressive fire used to support the final assault?	FIST
14. When contact was made, did platoon leader control the fire distribution of his elements?	PL
15. Did platoon leaders know the location of target reference points and preplanned fires?	PL
16. Was appropriate use made of target reference points and preplanned fires?	CO/PL/FIST
17. Was appropriate use made of indirect fire support?	CO/PL
17.' Was appropriate use made of indirect fire support?	FIST
18. Was communications security (COMSEC) maintained throughout the exercise?	CO/PL/FIST

1. Operations OrderEnemy Situation

_____ Did it include size, direction and type of enemy units?

Friendly Situation

Did it include:

_____ Battalion mission?

_____ Adjacent company missions?

Company Mission:

_____ Was mission stated?

Did it include: Who _____? What _____? Where _____? When _____? Why _____?

Execution

_____ Did the order include LD crossing time?

_____ Did the order identify the objective (N/A if no objective)

Were specific tasks assigned to:

_____ 1st Platoon?

_____ HAW Section?

_____ 2nd Platoon?

_____ GSR?

_____ 3rd Platoon?

_____ ENGR?

NOTE: Obtain phase lines, check points, preplanned fires, and target reference points from Commander's overlay.

Did scheme of maneuver include:

_____ Specific routes for each element?

_____ Phase lines?

_____ Check points?

_____ Overwatch assignments?

Fire Support Plan

Did it include:

_____ Preplanned fires?

_____ Target reference points?

_____ Priority of fires?

Command and Signal

Did it address:

____ CEOI?

____ Chain of command?

a) ^{••} RATING: Operations Order

very poor	unsatisfactory	borderline	satisfactory	very good

Confidence level in accuracy of evaluation:

0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%

2. Was the scheme of maneuver appropriate to the mission and terrain?

- . Did the CO specify general routes of advance for all elements? _____
- . Did the routes of advance selected by the CO for the elements take advantage of available cover and concealment? _____
- . Does maneuver scheme provide for continuous overwatch (when an overwatch movement technique is appropriate)? _____
- . Does the scheme of maneuver specify general location of overwatch positions? _____

a) **RATING:** Was the scheme of maneuver appropriate to the mission and terrain?

very poor	unsatisfactory	borderline	satisfactory	very good

Confidence level in accuracy of evaluation:

0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%

b) It was not possible to observe enough to make an evaluation: _____

Reason:

3. Was Fire Plan developed which would support the scheme of maneuver?

- . Were indirect fires planned to cover movement of the unit? _____
- . Did plan include use of smoke to mask elements moving through poorly covered areas? _____
- . Were preplanned fires designated on suspected enemy positions? _____
- . Were preplanned fires designated on, beyond, and on flanks of final objective (if appropriate)? _____

a) **RATING:** Was Fire Plan developed which would support the scheme of maneuver? (Consider also PL evaluator input regarding designation of target priorities for direct fire weapons).

very											very
poor		unsatisfactory		borderline		satisfactory					good

Confidence level in accuracy of rating:

0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%

b) No fire plan was developed: _____

c) It was not possible to observe enough to make an evaluation: _____

Reason:

3'. Were direct fire weapon systems assigned priority of targets?

- a. Before the mission did the CO, PL or SOP prescribe target engagement priority for TOWs? Yes _____ No _____

If yes, indicate source: (CO, PL, SOP) _____:

target priorities

(1) _____

(2) _____

(3) _____

- b. Before the engagement were target priorities designated for tanks (main gun)? Yes _____ No _____

If yes, indicate source: (CO, PL, SOP) _____;

target priorities

(1) _____

(2) _____

(3) _____

- c. Before the engagement were target priorities designated for DRAGONS? Yes _____ No _____

If yes, indicate source: (CO, PL, SOP) _____;

target priorities

(1) _____

(2) _____

(3) _____

NOTE: This information should be sought from crews in the assembly area before the exercise. The information obtained should be submitted to the CO evaluator(s) during or immediately after the exercise.

4. Was the LD crossed on time?

- a) Time for crossing LD specified in OPORD: _____
- b) Was the LD crossed on time? Yes _____ No _____
- c) If no, what was the amount of deviation in minutes? _____
- d) It was not possible to observe the crossing of the LD and/or no time was specified: _____

Reason:

5. Was movement covered by overwatch, when appropriate?

(Overwatch is normally appropriate when enemy contact is possible or expected.)

Observation #	Overwatch was appropriate and used (✓)	Overwatch was appropriate and not used (✓)	Remarks:
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			

a) RATING: Was movement covered by overwatch, when appropriate?

very poor unsatisfactory borderline satisfactory very good

|-----|-----|-----|-----|

Confidence level in accuracy of rating:

0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|

b) It was not possible to observe enough to make an evaluation: _____

Reason:

6. Did the bounding element minimize exposure during movement?

Bound #	Movement technique utilized: (T, TO, BO)	Was technique appropriate to situation? (Y, N)	Was smoke used to mask bound through open areas? (Y, N, Not Available)	Was available terrain and foliage used to mask movement? (Usually, Sometimes, Rarely)	Did the element use rapid movement across open areas? (Y, N)	Did they maintain vehicle dispersion? (Y, N)	Did they bound to covered/concealed positions? (Y, N, Not Available)
1							
2							
3							
4							
5							
6							
7							
8							
9							
10							

a) RATING: (Element rated _____) Did the bounding element minimize exposure during movement?

very poor unsatisfactory borderline satisfactory very good

Confidence level in accuracy of evaluation:

0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

b) It was not possible to observe enough to make an evaluation: _____

Reason:

7. Execution of overwatch

Observation #	Did overwatch provide:			Did overwatch position minimize vulnerability to enemy fire (i.e., through advantageous use of available cover + concealment)? (Y,N)	Was overwatch mounted or dismounted? (M,D)	Considering preceding points, were better overwatch positions available? (Y,N)
	a) maximum observation of the route of the bounding element? (Y,N)	b) observation of most potential enemy positions to which the bounding element would be exposed (Y,N)	c) effective fields of fire and responsive fire support? (Y,N)			
1						
2						
3						
4						
5						
6						
7						
8						
9						
10						

a) RATING: Did overwatch positions selected maximize observation ability?

very poor unsatisfactory borderline satisfactory very good

Confidence level in accuracy of evaluation:

0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

- | | | | | |
|-----------|----------------|------------|--------------|-----------|
| very poor | unsatisfactory | borderline | satisfactory | very good |
| 1 | 2 | 3 | 4 | 5 |

0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

- very poor unsatisfactory borderline satisfactory very good**

0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

- Reason:**

8. Was CO kept informed of friendly/enemy situation?

Reportable situations: reaching check point(s)/phase line(s) - Friendly
 suffering crew casualties - Friendly
 experiencing equipment failure - Friendly
 initiating action - Friendly
 receiving fire - Friendly + Enemy
 encountering obstacle (e.g., minefield) - Friendly + Enemy
 sighting enemy vehicle(s)/position(s)/signature(s) - Enemy

	(a)	(b)	(c)	(d)	(e)	(f)
TRANS- MISSION #	Type of Report (E/F)	(✓) if initiated by CO	CO received sufficient information (Y,N)	If not sufficient did CO pursue better in- formation (✓ = Yes)	If routine reports were not timely did CO pursue information? (✓ = Yes)	Remarks:
1						
2						
3						
4						
5						
6						
7						
8						
9						
10						
11						
12						
13						
14						
15						
16						
17						
18						
19						
20						

a) RATING: Was CO kept informed of friendly situation?

very poor unsatisfactory borderline satisfactory very good

Confidence level in accuracy of evaluation:

0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

b) RATING: Was CO kept informed of enemy situation?

very											very
poor		unsatisfactory		borderline		satisfactory					good

_____	_____	_____	_____	_____	_____	_____	_____	_____	_____	_____	_____
-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------

Confidence level in accuracy of evaluation:

0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
----	-----	-----	-----	-----	-----	-----	-----	-----	-----	------

_____	_____	_____	_____	_____	_____	_____	_____	_____	_____	_____
-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------

c) It was not possible to observe enough to make an evaluation: _____

Reason:

8'. Was CO kept informed of friendly/enemy situation?

Reportable situations:

reaching check point(s)/phase line(s)	- Friendly
suffering crew casualties	- Friendly
experiencing equipment failure	- Friendly
initiating action	- Friendly
receiving fire	- Friendly + Enemy
encountering obstacle (e.g., minefield)	- Friendly + Enemy
sighting enemy vehicle(s)/position(s)/signature(s)	- Enemy
displacement of fire units (FA, mortars)	- Friendly
changes to priority of fire (Bn/TF and Bde)	- Friendly

FIST {

Type of communication:

- Commo present (e.g. - routine time checks)
- Control measure (check points, phase lines, objectives)
- Friendly Status - SITREP - (actions, locations, casualties, equipment failure)
- Enemy Action - SPOTREP - (vehicles, units, signature, receipt of fire, obstacles)
- Other

Observation #

Commo Type: (✓)					Did PL inform CO? (Y,N)	Was any of the information <u>inaccurate</u> ? (Y,N)
Present	Control (friendly)	SITREP (friendly)	SPOTREP (enemy)	Other		
1						
2						
3						
4						
5						
6						
7						
8						
9						
10						
11						
12						
13						
14						
15						
16						
17						
18						
19						
20						

a) RATING: Was CO kept informed of friendly situation?

very											very
poor		unsatisfactory		borderline		satisfactory					good

Confidence level in accuracy of evaluation:

0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
----	-----	-----	-----	-----	-----	-----	-----	-----	-----	------

b) RATING: Was CO kept informed of enemy situation?

very											very
poor		unsatisfactory		borderline		satisfactory					good

Confidence level in accuracy of evaluation:

0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
----	-----	-----	-----	-----	-----	-----	-----	-----	-----	------

c) It was not possible to observe enough to make an evaluation: _____

Reason:

9. Was the FRAG Order complete and appropriate to the situation?

- . Did the circumstances of the engagement call for/require that a FRAG order be issued? Y _____ N _____ ? _____
- . Was a FRAG Order issued? Y _____ N _____
- . (If issued) did the FRAGO contain orders to all subunits _____, including fire support plan _____, and priorities of fires? _____
- . (If issued) did the FRAGO include control measures? _____
- . (If issued) did the FRAGO concentrate Company Team combat power, _____ and make maximum use of available direct and indirect fire? _____

a) RATING: Was the FRAG Order complete and appropriate to the situation?

very poor	unsatisfactory	borderline	satisfactory	very good

Confidence level in accuracy of evaluation:

0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

--	--	--	--	--	--	--	--	--	--	--	--

- b) (If issued) It was not possible to observe enough to make an evaluation: _____

Reason:

10. When engaged, did (a) engaged element, (b) supporting element take appropriate action?

Engagement #	Not observed or not applicable (✓)	Engaged by Direct Fire				Engaged by Indirect Fire			Action appropriate? (Y, N)
		return fire	deploy	report	request IF/DF fire	button up	move/seek cover	report	
1 Engaged Elem.									
1 Support Elem.									
2 Engaged Elem.									
2 Support Elem.									
3 Engaged Elem.									
3 Support Elem.									
4 Engaged Elem.									
4 Support Elem.									

- a) RATING: Did element engaged take appropriate action?

very poor unsatisfactory borderline satisfactory very good

Confidence level in accuracy of evaluation:

0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

- b) RATING: Did supporting element (if applicable and observed) take appropriate action?

Very poor unsatisfactory borderline satisfactory very good

Confidence level in accuracy of evaluation:

0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

- c) It was not possible to observe enough to make an evaluation: _____

Reason:

11. Did weapon systems move after engaging enemy targets?

Observation #	Weapon System moved after shooting (Y,N)	Instances where it would have been inappropriate to move (✓)
1		
2		
3		
4		
5		
6		
7		
8		
9		
10		

a) RATING: Did weapon system(s) move after engaging enemy target(s)?

very poor unsatisfactory borderline satisfactory very good

|-----|-----|-----|-----|

Confidence level in accuracy of evaluation:

0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|

b) It was not possible to observe enough to make an evaluation: _____

Reason:

12. After enemy contact is made, are elements controlled using appropriate fire and maneuver techniques?

- . Did the unit (elements) that made contact establish a base of fire (BOF) to provide direct and indirect fire support for maneuver?

Y _____ N _____

Remarks: _____

- . Did the maneuver unit at any time mask the overwatching fires or advance beyond the effective supporting range of the BOF? Y _____ N _____

Remarks: _____

- . Was fire and maneuver continuously coordinated between the BOF and maneuver elements? (i.e., were the elements in constant radio contact)?

Y _____ N _____

Remarks: _____

- . Were enemy positions within reach suppressed or destroyed so that the maneuvering element(s) could continue to advance? Y _____ N _____

Remarks: _____

- a) **RATING:** After enemy contact is made, are elements controlled using appropriate fire and maneuver techniques?

very												very
poor		unsatisfactory		borderline		satisfactory						good

Confidence level in accuracy of evaluation:

0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%

- b) It was not possible to observe enough to make an evaluation: _____

Reason: _____

13. Was suppressive fire used to support the final assault?

. Did the leader observed have assets available:

Non-organic direct fire? _____ Indirect fire? _____

. Was suppressive fire used on objective/enemy positions prior to and during assault:

Non-organic direct fire? _____ Indirect fire? _____

. Was suppressive fire shifted to the rear and flanks of the enemy during the assault?

Yes _____; No _____; Wasn't Appropriate _____; Could not determine _____.

a) RATING: Was suppressive fire used to support the final assault (given the assets that were available)? Note: Incorporate input from FIST evaluator.

very											very
poor		unsatisfactory		borderline		satisfactory					good

Confidence level in accuracy of evaluation:

0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%

b) It was not possible to observe enough to make an evaluation: _____

Reason:

FIST

13'. Was suppressive fire used to support the final assault?

. Were IF assets available to CO _____, PL _____ at the time of the assault?

. Did FIST chief utilize suppressive IF on objective/enemy positions in support of the assault? (Y,N) _____

HE? _____

Smoke? _____

a) **RATING:** Was suppressive IF used to support the final assault (given the assets that were available)?

very												very
poor		unsatisfactory		borderline		satisfactory						good

Confidence level in accuracy of evaluation:

0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%

b) It was not possible to observe enough to make an evaluation: _____

Reason:

NOTE: Submit this information to CO evaluator immediately after the exercise for his overall evaluation of this item.

14. When contact was made, did platoon leader control the fire distribution of his elements?

. At time of contact, did PL know where his elements were? Y___ N___

. Did the PL assign specific targets or sectors to his squads or sections? Y___ N___

. Were most engagements "one-on-one" type of situations? Y___ N___

. Did platoon have SOP for fire distribution on contact? Y___ N___

a) RATING: When contact was made, did platoon leader control the fire distribution of his elements?

very poor	unsatisfactory	borderline	satisfactory	very good

Confidence level in accuracy of evaluation:

0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%

b) It was not possible to observe enough to make an evaluation: _____

Reason:

15. Did platoon leaders know the location of target reference points and preplanned fires?

a) Yes _____ No _____

Conditions for observation:

- Platoon leader calls for TRP/PF fires.
- After exercise is completed ask platoon leader for TRP/PF locations, or look at PL's overlay.

b) It was not possible to observe enough to make an evaluation: _____

Reason:

16. Was appropriate use made of target reference points (TRP) and preplanned fires (PF)?

. Were TRPs selected/requested by CO? Y ☐ N ☐
 PL? Y ☐ N ☐

. Were TRPs located near distinguishable landmarks, such as:

road junctions? ☐
 lone tree or rock formation? ☐
 hill top? ☐
 tree line? ☐
 Other (specify)

. Was there generally one TRP per 1000 meters (i.e., each grid square)? ☐

. Were TRP/PFs visible from successive friendly positions? ☐

. Were TRPs used to control maneuver or fire distribution? ☐

. When TRP/PF was near a target, was it used? Y ☐ N ☐ (Keep a tally if there are multiple observations.)

a) RATING: Was appropriate use made of TRPs and PFs?

very											very
poor		unsatisfactory		borderline		satisfactory					good

Confidence level in accuracy of evaluation:

0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%

b) It was not possible to observe enough to make an evaluation: ☐

Reason:

17. Was appropriate use made of indirect fire (IF) support?

- . Were good opportunities to use IF overlooked? _____ How many times? _____
- . Were IF missions used when not needed? _____ How many times? _____
- . Were appropriate means/assets requested for IF missions? _____
- . Was fire support coordinated efficiently? _____

Obtain the following from the FIST evaluator after the exercise:

- . Number of smoke missions available to the unit _____
- . Number of smoke missions called by the unit: _____
- . Number of HE missions available to the unit: _____
- . Number of ARTY missions called by the unit: _____
- . Number of mortar missions called by the unit: _____

- a) **RATING:** Was appropriate use made of indirect fire (considering the IF assets available, and used)?

very											very
poor		unsatisfactory		borderline		satisfactory					good

Confidence level in accuracy of evaluation:

0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%

- b) It was not possible to observe enough to make an evaluation: _____

Reason:

17. Was appropriate use made of indirect fire (IF) support?

- . How many smoke missions were available to the unit? _____
- . How many smoke missions were called by the unit? _____
- . How many HE missions were available to the unit? _____
- . How many ARTY missions were called by the unit? _____
- . How many mortar missions were called by the unit? _____

NOTE: The above information should be shared with the CO and PL evaluators immediately after the exercise.

- . Were good opportunities to use IF overlooked? _____ How many times? _____
- . Were IF missions used when not needed? _____ How many times? _____
- . Were appropriate means/assets requested for IF missions? _____
- . Was fire support coordinated efficiently? _____
- a) RATING: Was appropriate use made of indirect fire (considering the IF assets available and used)?

very											very
poor		unsatisfactory		borderline		satisfactory					good

Confidence level in accuracy of evaluation:

0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%

- b) It was not possible to observe enough to make an evaluation: _____

Reason:

18. Was communications security (COMSEC) maintained throughout the exercise?

- . Were violations of COMSEC observed during the exercise (e.g., transmission of unit coordinates or names of individuals or organizations in clear)?

Y _____ N _____

- . If observed, how many COMSEC violations occurred and of what type?

Number observed: _____

Type(s): _____

- a) **RATING:** Was communications security (COMSEC) maintained throughout the exercise?

very												very
poor		unsatisfactory		borderline		satisfactory						good

Confidence level in accuracy of evaluation:

0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%

- b) It was not possible to observe enough to make an evaluation: _____

Reason:

APPENDIX B: Defensive Rating Scale for Previous Dunn-Kempf Study:
"An Experimental Evaluation of a Tactical Game for Company
Level Training"

I. OPERATIONS ORDER

Enemy Situation

_____ Did it include size, direction and type of enemy units?

_____ Was the tac air situation mentioned?

Friendly Situation: Did it include:

_____ Battalion mission

_____ Adjacent companies mission

Company Mission: Was it given?

Execution:

Were specific tasks assigned to:

_____ 1st Plat

_____ 2nd Plat

_____ 3d Plat

_____ HAW Sect

_____ Engrs

_____ Was a completion time designated?

_____ Was someone assigned the passage point?

_____ Were sectors of responsibility designated?

Fire Support Plan

Did it include:

_____ Preplanned Fires

_____ Registered Fires

_____ Were fires called to battalion FDC?

_____ Final Protective Fires

Barrier Plan: If a barrier plan existed was it mentioned?

Coordinating Instructions:

_____ Was Passage Point recognition signal mentioned?

Service Support:

_____ Was location of Company trains given?

Command and Signal

Did it include location of

_____ CO

_____ FO

_____ Commo track

_____ Was CEOI addressed?

II. EXECUTION OF OPERATIONS ORDER

(Compare CO's sketch of his intended defense plan with actual set up.)

Were platoons located in positions designated by CO?

Were intended platoon sectors of responsibility covered?

_____ Did someone man the passage point?
_____ Did Platoon Leaders know the location of registered and
preplanned fires in their sectors?

III EVALUATION OF DEFENSIVE PLAN

_____ Did Platoon positions selected by the CO afford/provide:
_____ Cover
_____ Concealment
_____ Use of reverse slope advantage
_____ Long range vision
_____ Long range fields of fire
_____ Mutual supportability with interlocking fire between platoons
_____ Concentration of massed direct fire on most likely routes
_____ of enemy advance
_____ Covered and trafficable routes of withdrawal
_____ Primary and Supplementary platoon positions

_____ Did individual HAW positions afford/provide:
_____ Cover
_____ Concealment from air observation; i.e., woods or hide
_____ position behind firing position
_____ Alternate defilade firing positions
_____ Long range fields of fire (out to 3000m)
_____ Enough distance from other firing vehicles to prevent
_____ simultaneous suppression by a single artillery mission

_____ Did individual vehicle positions afford/provide:
_____ Cover
_____ Concealment from air observation, i.e., in woods or hide
_____ position behind firing position
_____ Long range field of fire (at least up to 1500m)
_____ Primary and alternate firing positions (either natural or
_____ man-made) within 50m.

_____ Was the passage point covered by fire.

_____ Indirect fires
_____ Were preplanned fires located along all likely avenues of
_____ enemy approach?
_____ Were registered fires located along most likely enemy
_____ avenues of approach?
_____ Were registered fires located so as to be easily adjusted
_____ from?

_____ Barrier Plan:
_____ Were barriers placed in areas not easily bypassed?
_____ So as to channelize the enemy or to deny him avenues of
_____ approach?

IV. CONDUCT OF THE DEFENSE

_____ Did CO shift platoons to meet major enemy threat?

Passage of Lines

_____ Did friendly vehicles man the passage point until all passing vehicles had passed?

Did HAWS:

_____ Open fire at maximum range (2000-3000m)?

_____ Shoot and move to avoid being suppressed?

_____ Coordinate fire with tank direct fire suppression so as to minimize vulnerability of HAW to enemy counter-fire?

_____ Give priority of fires to most dangerous enemy weapons?

_____ *Did DRAGON contribute to the defensive effort?

_____ Were initial engagements conducted so as to minimize divulgence of friendly strength and positions?

Tank fire techniques

_____ Was direct fire used to suppress the most lethal enemy weapons?

_____ Once enemy vehicles had moved within effective range of friendly positions, did friendly vehicles shoot and move to alternate positions to reduce their vulnerability?

Reports

_____ Was the CO kept informed of friendly situation?

_____ Was the CO kept informed of enemy situation?

_____ *Was clearance of the passage point reported?

_____ Was maximum use made of registration points?

_____ Were adjustments quick and accurate?

_____ Was ARTY fire used to suppress most dangerous enemy weapons?

*These items were graded as dichotomous, with values: OK=4, not OK=2.

Table B-1
Dunn-Kempf Measures

High Gain Items:

1. Did HAWS give priority of fires to most dangerous enemy weapons?
2. Once enemy vehicles had moved within effective range of friendly positions, did friendly vehicles shoot and move to alternate positions to reduce their vulnerability?
3. Did HAWS coordinate fire with tank direct fire suppression so as to minimize vulnerability of HAW to enemy counter-fire?
4. Did individual HAW positions afford/provide alternate defilade firing positions?
5. Did individual vehicle positions afford/provide primary and alternate firing positions (either natural or man-made) within 50m.
6. Was the CO kept informed of enemy situation?
7. Was direct fire used to suppress the most lethal enemy weapons?
8. Were ARTY adjustments quick and accurate?
9. Did platoon positions selected by the CO afford/provide cover?
10. Was the CO kept informed of friendly situation?
11. Did CO shift plts to meet the major enemy threat?

Moderate Gain Items:

12. Did platoon positions selected by the CO afford/provide use of reverse slope advantage?
13. Were registered fires located so as to be easily adjusted from?
14. Did platoon positions selected by the CO afford/provide long range fields of fire?
15. Did platoon positions selected by the CO afford/provide mutual supportability with interlocking fire between platoons?
16. Did individual HAW positions afford/provide cover?

Table B-1(Con't.)
Dunn-Kempf Measures

Moderate Gain Items: (Con't.)

17. Did HAWS open fire at maximum range (2000-3000m)?
18. Did DRAGON contribute to the defensive effort?
19. Was ARTY used to suppress the most dangerous enemy wpns?
20. Did individual HAW positions afford/provide concealment from air observation; i.e., woods or hide position behind firing position?
21. Did platoon positions selected by the CO afford/provide concentration of massed direct fire on most likely routes of enemy advance?
22. Did individual vehicle positions afford/provide cover?
23. Were preplanned fires located along all likely avenues of approach?
24. Were registered fires located along most likely enemy avenues of approach.
25. Did HAWS shoot and move to avoid being suppressed?
26. Was maximum use made of registration points?
27. Did platoon leaders know the location of registered and preplanned fires in their sectors?
28. Did platoon positions selected by the CO afford/provide concealment?
29. Did individual vehicle positions afford/provide concealment from air observation, i.e., in woods or hide position behind firing position?
30. Did platoon positions selected by the CO afford/provide long range vision?

Low Gain Items:

31. Did individual vehicle positions afford/provide long range field of fire (at least up to 1500m)?
32. Were intended platoon sectors of responsibility covered?

Table B-1(Con't.)

Dunn-Kempf Measures

Low Gain Items: (Con't.)

33. Did individual HAW positions afford/provide enough distance from other firing vehicles to prevent simultaneous suppression by a single artillery mission?
34. Was the passage point covered by fire?
35. Did platoon positions selected by the CO afford/provide primary and supplementary platoon positions?
36. Were initial engagements conducted so as to minimize divulgence of friendly strength and positions?
37. Were barriers placed so as to channelize the enemy or to deny him avenues of approach?
38. Did platoon positions selected by the CO afford/provide covered and trafficable routes of withdrawal?
39. Were barriers placed in areas not easily bypassed?
40. Were platoons located in positions designated by CO?

Appendix C: Evolution of Original Dunn-Kempf Measures to Finalized Performance Evaluation Instrument

The purpose of this appendix is to trace the evolution of the original Dunn-Kempf measures developed during the earlier ARI study (1977), to the instrument which has been developed for use in a transfer of training experiment and which is considered applicable to both Dunn-Kempf and REALTRAIN exercises. As was indicated in the main body of the report, a number of measures were not usable because they related exclusively to the defense mission and a realistically-opposed company-level defense is rarely if ever conducted in a field exercise. Additionally, the ten "low-gain" measures from the original instrument were deleted from consideration on the basis of having low potential for transfer. Each of the remaining 30 measures from Table B-1, the original Dunn-Kempf measures, will be discussed below. To the extent that measures have been retained or revised, they are referenced to items in the current instrument presented in Appendix A.

1. Did HAW give priority of fires to most dangerous enemy weapons? This item was not considered collectable in its original form in REALTRAIN exercises because: It measured performance of the HAW crew chiefs who did not participate in Dunn-Kempf exercises; it depended on availability of multiple targets of varying lethality; and it required too much inference on the part of the evaluator. Additionally, the real time aspect of REALTRAIN would have made data collection difficult. However, the concept of target prioritization is retained for HAW (and other direct-fire weapons) in current item number 3, "Were direct-fire weapon systems assigned priority of targets?" This measure reflects on the PL and CO and is incorporated in the overall evaluation of the fire plan.

2. Once enemy vehicles had moved within effective range of friendly position, did friendly vehicles shoot and move to alternate positions to reduce their vulnerability? This item was eliminated as relating only to defensive missions. The tactical concept of "shoot and move" is retained in current item number 11, "Did weapon systems move after engaging enemy targets?"
3. Did HAWS coordinate fire with tank direct-fire suppression so as to minimize vulnerability of HAW to enemy counter-fire? This item was eliminated on the basis of being difficult to observe and therefore requiring undue inference on the part of the data collector, and also on the basis that it would evaluate players not involved in Dunn-Kempf exercises so that occurrence of transfer would be equivocal.
4. Did individual HAW positions afford/provide alternate defilade firing positions? This item was eliminated as applicable only to defensive missions. The concept of taking-up defilade firing positions is retained in item number 7c, "Did the overwatch positions selected minimize vulnerability to enemy detection and fire?"
5. Did individual vehicle positions afford/provide primary and alternate firing positions (either natural or man-made) within 50 m? This item was eliminated as applying only to defensive missions.
6. Was the CO kept informed of enemy situations? This item was retained as is in current measure number 8. In the new instrument it is monitored from both the CO and PL positions and rated separately. Also, for ease of data collection, commo pertaining to both friendly and enemy situations are combined into a single data collection form, although the measures are rated separately for enemy and friendly situations.

7. Was direct fire used to suppress the most lethal enemy weapons? The mechanisms for playing and reporting suppression in REALTRAIN are not well developed nor consistently used. Therefore, this measure was eliminated as being too difficult to observe and interpret in the field exercises. However, the concept of employing suppressive means is retained in current items 12 and 13.
8. Were ARTY adjustments quick and accurate? This item was eliminated as not being meaningfully measurable in the field. Both the speed and accuracy of ARTY adjustments in REALTRAIN are partially a function of personnel involved in the administration and control of the exercises. A second and equally compelling reason for discarding the measure was that adjustments from original fire requests were almost never made during the REALTRAIN exercises observed.
9. Did platoon positions selected by the CO afford/provide cover? The concept of this item, which pertains to the defense in its original form, was revised for an offensive mission to deal with selection of maneuver routes by the CO. This now constitutes a sub-element of item number 2, "Was the scheme of maneuver appropriate to the mission and terrain?".
10. Was the CO kept informed of friendly situations? This item was retained as it is in current measure number 8. (See also explanation with item number 6.)
11. Did CO shift platoons to meet the major enemy threat? This item was revised to fit the offensive mission and now focuses on the use of fire and maneuver techniques in current measure number 12.
12. Did platoon positions selected by the CO afford/provide use of reverse

slope advantage? This item was eliminated as pertaining only to a defensive mission.

13. Were registered fires located so as to be easily adjusted from? Registered fires are rarely available in an offensive mission. Therefore, this item was revised to focus on the selection and use of target reference points (TRP) in current measure number 16.

14. Did platoon positions selected by the CO afford/provide long-range fields of fire? This item, which pertains essentially to a defense, was revised to relate to the offensive mission and now deals with the selection of overwatch for maximum observation and effective fire support (current items 7a and 7b).

15. Did platoon positions selected by the CO afford/provide mutual supportability with interlocking fire between platoons? This item was eliminated as relating only to a defensive mission. The basic consideration of mutual supportability is retained in items dealing with execution of overwatch and use of fire and maneuver techniques (current measures 7 and 12).

16. Did individual HAW positions afford/provide cover? This item was revised for the offensive mission and no longer differentiates among types of vehicle or weapon systems. (See also item 22.)

17. Did HAWs open fire at maximum range (2000-3000 m)? In an offensive scenario the HAW engagement range would be situationally dependent and would not necessarily be an indicator of good or poor performance. Therefore, this item was eliminated.

18. Did DRAGON contribute to the defensive effort? This item pertains to the defensive mission and was eliminated as a rating item. However, casualties by

weapon type (DRAGON, tank, ARTY, etc.) can be systematically recorded during both Dunn-Kempf and REALTRAIN exercises.

19. Was ARTY used to suppress the most dangerous enemy weapons? This item was eliminated in its original form. Dangerousness of a weapon is a function of its range, lethality and relative position (and may be dangerous to some targets and not others). Therefore, this item was judged very difficult to observe and interpret in a field exercise. Also, in REALTRAIN, ARTY suppression is usually effected upon a position rather than a specific (type of) weapon system. However, the use of indirect fire suppression is still evaluated in current measure number 13, "Was suppressive fire used to support the final assault?".

20. Did individual HAW positions afford/provide concealment from air observation, i.e., woods or hide position behind firing position? This item was eliminated as pertaining primarily to the defense. Additionally, the threat of air attack is not played in REALTRAIN and therefore may not be a consideration in selection of overwatch positions.

21. Did platoon positions selected by the CO afford/provide concentration of massed direct fire on most likely routes of enemy advance? The factors of platoon positioning and enemy routes of advance define this as a defensive measure, and as such it was eliminated. The consideration of massing of fire power is retained in current item number 9, "Was the FRAG Order complete and appropriate to the situation?".

22. Did individual vehicle positions afford/provide cover? This item was revised for the offensive mission to apply to use of cover during movement and at the ends of bounds (for both bounding and overwatch elements). The concept of use of cover is contained in items 6 and 7c. (This explanation

applies also to item 16.)

23. Were preplanned fires located along all likely avenues of approach?

This item was revised for the offensive mission to cover preplanned fires in advance of friendly approach. Additionally, it was expanded to cover other aspects of fire planning. This is the subject of current measure number 3, "Was fire plan developed which would support the scheme of maneuver?".

24. Were registered fires located along most likely enemy avenues of approach? This item was revised to consider target reference points in the offense and was incorporated in current item number 3.

25. Did HAWs shoot and move to avoid being suppressed? The concept of shoot and move is retained in current item number 11, "Did weapon system move after engaging enemy targets?". However, it no longer differentiates by weapon system.

26. Was maximum use made of registration points? This item was revised for the offensive mission to consider use of target reference points and to incorporate use of preplanned fires as well. Also, the word "maximum" was changed to "appropriate". The revised item is number 16 in the current instrument.

27. Did platoon leaders know the location of registered and preplanned fires in their sectors? This item was retained nearly in its original form. By way of adaptation to the offensive mission reference to registered fires was changed to target reference points and reference to sectors was dropped. This became current item number 15.

28. Did platoon positions selected by the CO afford/provide concealment?

This item was revised for the offensive mission to consider concealment for

routes of advancement rather than defensive positions. This concept is included in current item number 2, "Was the scheme of maneuver appropriate to the mission and terrain?".

29. Did individual vehicle positions afford/provide concealment from air observation, i.e., in woods or hide position behind firing positions? This item was eliminated as pertaining primarily to the defense. (See also item 20.)

30. Did platoon positions selected by the CO afford/provide long-range vision? This item was revised for the offense to consider the selection of successive overwatch positions. Further, the focus of the measure was shifted to the platoon leader (PL) because functionally in the offense overwatch positions are selected at this level and below. The concept of the item is retained in current measure number 7a, "Did overwatch positions selected maximize observation ability?".

ADDITIONAL MEASURES

Because a number of measures from those originally developed for Dunn-Kempf were eliminated, an analysis of offensive operations was undertaken to look for areas or categories of tactics for which new measures needed to be developed for evaluation to cover the entire scope of an exercise. Current Army doctrine, in the form of the Combined Arms ARTEP (71-2) and various How to Fight Manuals and Training Circulars, was reviewed for the identification and development of additional measures. The origin of new measures, and the elaboration of previous measures which was not described above, is discussed below. This discussion is keyed to the current measure numbers as presented in Appendix A.

The Operation Order (#1) was monitored during the previous Dunn-Kempf research, but it was not a rated item. However, the format for evaluating the operations order was adopted from the original Dunn-Kempf instrument nearly in its entirety. Minor modifications reflect the differences between defensive and offensive scenarios or the elimination of components not normally involved in or relevant to REALTRAIN exercises.

Crossing the line of departure (LD) on time (#4) is a training/evaluation standard specified in ARTEP 71-2 for a deliberate attack mission.

The use of overwatch to cover movement (5 & 7) is an area of tactics which is important to the offensive mission. Its use and techniques are emphasized in the ARTEP. For purposes of the evaluation instrument overwatch was considered from four aspects: Whether it was used when it was appropriate, and if used, the extent to which it provided good visibility, good ability to react, and good cover and concealment.

The use of FRAG Orders (#9) is not exclusive to the offense, but more common in offensive situations, particularly as the mission shifts from movement-to-contact to hasty attack. The appropriateness of a FRAG Order to the combat situation encountered is a key indicator of a leader's ability to assess and respond to his circumstances and requirements. The use and clarity of the FRAG Order is represented in ARTEP 71-2.

Appropriate reaction to engagement (#10), differs between the defensive and offensive missions and within the offensive mission depending on whether the observed element is engaged (#10a), or in support of a unit which is engaged (#10b). Therefore, this measure was developed to evaluate leaders' reactions at such critical times during the exercise. Prescribed reactions,

depending on the role of the element, are specified in How-to-Fight Manuals.

Fire and maneuver techniques represent an area of tactics which is specific to offensive missions. The control of these complex functions (distribution of fire while maneuvering elements) is both difficult to execute and difficult to judge, but it is an important aspect of leader performance which warrants evaluation.

The control of fire distribution (#14) is basic to all types of tactical operations and, to the extent practicable, should be handled in a predetermined systematic way. One way of looking at fire control is that it is a type of resource management. It is both important that the task be accomplished and that valuable resources be conserved. There are many options or techniques available for fire control, depending on the situations encountered. The important thing is that there is provision for fire control and thus it is put into effect when needed.

Because existing and revised measures looked at rather specific aspects of indirect fire utilization, it was decided to take a more general or summary view of the use of indirect fire. It has often been subjectively reported that battle simulations are useful for introducing awareness of the utility and capabilities of indirect fire. (This is probably because until fairly recently, indirect fire could only very crudely and arbitrarily be represented in field exercises.) Therefore, a measure was developed to evaluate overall utilization of indirect fire support (#17) as an important aspect of the mastery of tactics and as a potentially fruitful area for the measure of transfer.

The last additional area tapped in the development of the current instru-

ment deals with communications security (COMSEC). The concept of COMSEC is basic to all types of tactical operations. The primary motivation to include this item (#18) in evaluation was the large number of security violations noted during REALTRAIN exercises at Baumholder. Additionally, it is possible to play Dunn-Kempf requiring players to utilize formal reporting channels and procedures, and the preceding Dunn-Kempf study found greater improvement in communication-related items for groups in which formal commo procedures were imposed.

A final notation is relevant regarding the development of the current instrument. The original Dunn-Kempf measures were designed before the introduction of the fire-support team (FIST) concept throughout the Army. While this fact didn't generally change the intent of measures dealing with the utilization of indirect fire (IF), it did serve to change the focus of evaluation from the forward observer (FO) to the FIST Chief and also affected some of the components or procedures that are observed in rating IF-related measures.

Attachment

Research Plan
for Measuring Transfer of Training
Between Tactical Training Systems

Submitted to:
ARI Field Unit, USAREUR
24 September 1979

HumRRO
Human Resources Research Organization
300 N. Washington Street
Alexandria, Virginia 22314

TABLE OF CONTENTS

	<u>Page</u>
BACKGROUND	D-5
Transfer of Training	D-5
Training Effectiveness of Battle Simulations	D-6
PURPOSE	D-7
METHOD	D-7
Experimental Design	D-8
PROCEDURES	D-9
1. Identify participating units	D-9
2. Train REALTRAIN evaluators	D-11
3. Collect reliability data	D-11
4. Analyze data for reliability	D-12
5. Design REALTRAIN exercises	D-12
6. Conduct REALTRAIN pretest	D-12
7. Orient Dunn-Kempf trainer/controllers	D-13
8. Train Dunn-Kempf evaluators	D-13
10. Conduct REALTRAIN posttest	D-14
11. Analyze data for transfer of training	D-14
12. Prepare final report	D-18
APPENDIX A: Dunn-Kempf/REALTRAIN Performance Evaluation Form	D-21
APPENDIX B: Exercise Record Form	D-57

TABLE OF CONTENTS (Continued)

	<u>Page</u>
Figure 1. Experimental Design for Transfer of Tactical Training Experiment	D-9
Figure 2. Procedural Approach to a Dunn-Kempf/REALTRAIN Transfer of Training Experiment	D-10
Figure 3. Item Analysis of Performance Differentials Between the Experimental and Control Groups	D-15
Figure 4. Two-Way ANOVA Model	D-17
TABLE 1. Research Plan Support Requirements	D-19

Research Plan
for Measuring Transfer of Training
Between Tactical Training Systems

BACKGROUND

To train military leaders in tactical command skills the U.S. Army has developed a number of battle simulation and wargaming techniques. These techniques range in complexity from small playing boards with manually controlled playing pieces involving a few players, to large-scale computer controlled exercises for battalion and higher levels of play. Battle simulations and war games share common utilization characteristics and training advantages. They provide unit commanders, their staffs, and subordinate leaders the opportunity to practice critical command and control skills at greatly reduced cost and resource levels when compared to traditional field exercises.

These skills are developed and practiced in a simulated functional context in which players experience their tactical roles during the course of a battle. They have to make decisions and exercise their command functions in such critical areas as teamwork, tactics, maneuver, terrain utilization, and weapons system employment and lethality, in response to constantly changing battlefield events.

Battle simulations have a high degree of face or content validity and usually provide a high degree of enthusiasm and motivation among player personnel. However, a very real training effectiveness question remains to be answered. Namely, to what extent are command related skills learned or enhanced in these exercises and how well, if at all, do these skills transfer to the operational field environment? As these techniques become integrated into unit readiness training, the question of transfer of training effects increases in importance.

Transfer of Training

Transfer of training refers to the extent to which skills learned in one type of training setting will produce an effect on performance in another setting. The types of skills subject to transfer can range from simple psychomotor responses to the application of complex principles. The extent to which performance is enhanced or degraded in a "criterion environment" determines the degree of training transfer.

There are three types of transfer effect: Positive transfer relates to cases where a response or skill learned in one situation facilitates performance in a subsequent, or criterion situation. Zero transfer refers to instances when a previously learned skill has no measurable effect on performance in the criterion environment. Negative transfer occurs when a previously learned skill has a detrimental effect on subsequent performance. Functionally, transfer can occur over a continuum

from positive to negative transfer effects depending on such factors as the similarity of stimuli or cues within the two environments, and the appropriateness of the learned responses to the requirements of the criterion environment.

Training Effectiveness of Battle Simulations

Because of increased reliance on battle simulations for the training of tactics and leadership skills, there is a great deal of interest in determining their effectiveness. Studies of the training effectiveness of several of the current generation of battle simulations have been conducted to investigate what types of skills and principles are learned through experience with these simulations.^{1,2,3} The studies have typically utilized some combination of objective and subjective performance measures and surveys of participating players to evaluate learning and performance within these settings. However, to date there has been no attempt to measure the effects of training in one battle simulation on performance in a field tactical training environment. Over the last two years the Army Research Institute Field Unit in Heidelberg has been engaged in a research program aimed at developing a methodology to measure transfer of training from the Dunn-Kempf board game to the REAL-TRAIN field training environment.

The initial effort in this research program involved an experimental evaluation of the Dunn-Kempf game.⁴ The evaluation research resulted in the development of both subjective and objective performance measures which were used to evaluate training progress and alternative training strategies. The results showed that in terms of what is learned in the game, improvement was most striking in three skill domains: (a) relative priority assigned to high-threat targets; (b) coordination among team members; and (c) shifting of forces as the battle develops.

The initial study was followed by the development of a transfer of

¹ I. T. Kaplan and H. F. Barber, "Evaluation of a Computer-Assisted Battle Simulation: CAMMS versus a CPX", U. S. Army Research Institute Technical Paper 355, April, 1979.

² H. F. Barber and I. T. Kaplan, "Battalion Command Group Performance in Simulated Combat," U. S. Army Research Institute Technical Paper 353, March, 1979.

³ E. E. Miller and J. M. Bachta, "An Experimental Evaluation of a Tactical Game for Company Level Training," U. S. Army Research Institute Draft Technical Report, January, 1978.

⁴ Ibid.

training methodology designed to measure transfer between the Dunn-Kempf battle simulation and REALTRAIN tactical field exercises. Dunn-Kempf was selected because it is widely disseminated and accessible in USAREUR, and therefore findings relative to transfer would have immediate training implications. Also, because Dunn-Kempf is designed to simulate company and platoon level operations it was considered feasible to duplicate these exercises in a field environment for the measure of transfer (all other Army-wide battle simulations focus at battalion level or higher). REALTRAIN field exercises were selected as a criterion environment based on the assessment that REALTRAIN is the closest analog to Dunn-Kempf among the various field exercises and it is also the most realistic simulation of a combat environment.

The major thrust of the methodological development was the construction of a performance evaluation instrument which could be used in both Dunn-Kempf and REALTRAIN exercises. The instrument covers a broad range of tactical skills and activities and focuses on both specific and general aspects of tactical performance. Performance measurement is achieved by using trained military evaluators who assign ratings for each of the applicable measures. The measures themselves are based on those used in the original Dunn-Kempf evaluation and have been refined during two phases of field application and revision. The performance evaluation instrument in its final form is at Appendix A. Documentations of the methodological development and instrument development has been submitted separately in the project final report, "Development of a Methodology for Measuring Transfer of Training Effects for Tactical Training Systems", on 24 September 1979. It was originally planned to assess the reliability of the performance evaluation instrument as part of the methodological development. However, because of the unavailability of field support, it was not possible to perform this assessment. Therefore, this research plan includes a requirement to collect instrument reliability data prior to the transfer of training study.

PURPOSE

The purpose of the present paper is to present a research plan for measurement and assessment of transfer of training between a battle simulation (Dunn-Kempf) and a tactical field exercise (REALTRAIN).

METHOD

The transfer of training experiment will require the evaluation and comparison, in REALTRAIN field exercises, of experimental units that have participated in structured Dunn-Kempf training and control units that have not had recent or systematic Dunn-Kempf training. Performance will be evaluated in specific tactical skill domains for both Dunn-Kempf and REALTRAIN field exercises by teams of trained military data collectors using a performance evaluation instrument developed for this purpose. Performance measures for experimental and control groups will be compared to determine to what extent Dunn-Kempf training has an effect on subsequent

field performance. In addition to the performance evaluation, objective outcome data, in the form of casualty assessment records, will be collected for each exercise and between-group comparisons will be made.

As an initial part of this effort, an investigation of performance instrument reliability is necessary. The investigation of reliability is concerned with assessing whether differences in ratings reflect actual performance differentials or whether they may result from instability of the measures used. The confidence with which the occurrence of transfer can be concluded is dependent on the known reliability of the measurement instrument. As discussed below, the reliability analysis will be accomplished by comparing independent ratings of evaluators observing the same performance.

Experimental Design

The transfer of training experiment will employ a non-equivalent control group design.⁵ This type of design is necessary because it will not be possible to randomly assign individuals to experimental or control groups (since they will operate in their unit configuration) and because there will be no basis for assuming that the groups are equally proficient prior to participation in the experiment. The experimental design, represented graphically in Figure 1, requires a pretest and a posttest for both experimental and control units with intervening Dunn-Kempf training for the experimental units. The experimental treatment will be randomly assigned to one of the paired units.

This pretest-posttest design will be repeated for a series of experimental and control groups to increase the reliability and generalizeability of the results. The experiment should involve a minimum of three such iterations and as many more as project resources can accommodate. This approach should afford flexibility in scheduling, and is also the most realistic in view of the scale of the effort and the availability of supporting units and exercises. It diminishes the considerable coordination and support effect that would be required to conduct the experiment concurrently with multiple experimental and control units.

The use of the pretests and control groups affords a greater potential for inference relative to transfer effects than either single group of posttest only designs. Use of the pretest has the additional advantage of assuring familiarity with REALTRAIN so that posttest performance is not unduly depressed by the complexity of an engagement simulation exercise as was found in a previous study.⁶

⁵ D. T. Campbell and J.C. Stanley, Experimental and Quasi-Experimental Designs for Research, Rand McNally & Company, Chicago, 1972.

⁶ R. T. Root, J. Hayes, et al., "Project EFFTRAIN: Field Test of Techniques for Tactical Training of Junior Leaders in Infantry Units," U. S. Army Research Institute Draft Technical Report, November, 1975.

Figure 1. Experimental Design for Transfer of Tactical Training Experiment

	REALTRAIN PRETEST	DUNN-KEMPF TRAINING	REALTRAIN POSTTEST
Experimental Group	X	X	X
Control Group	X	—	X

A final aspect of the design is that different evaluator teams will be used to evaluate REALTRAIN and Dunn-Kempf exercises. The REALTRAIN (criterion) evaluators will be "blind" with respect to the identity of experimental and control units in order to avoid possible experimenter bias effects (e.g., "halo" effect). As a minimum requirement, the same evaluator team should rate the performance of both experimental and control groups for each iteration, and the use of the same team for pretest and posttest would contribute further to overall reliability of measurement.

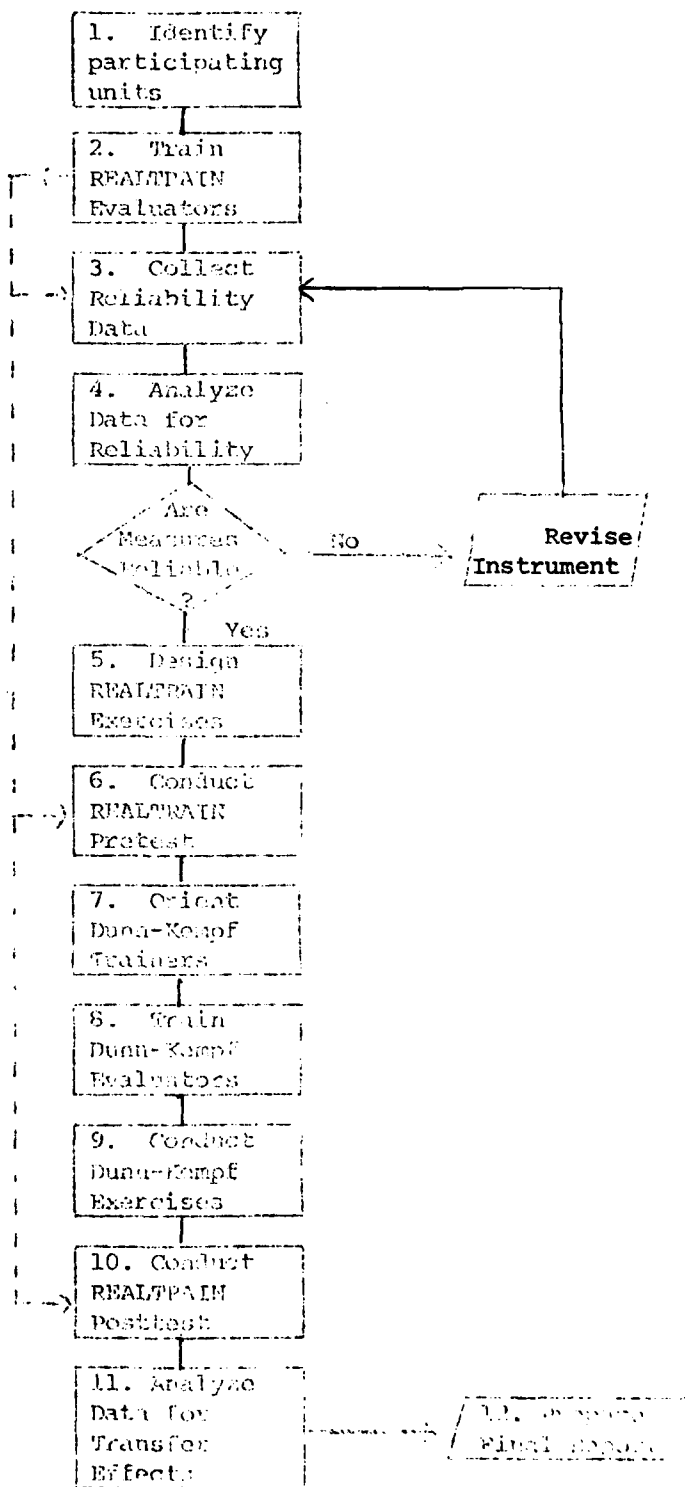
PROCEDURES

Figure 2 depicts the sequence of subtasks required to accomplish the research. Subtask numbers correspond to the procedural descriptions in the following paragraphs.

1. Identify participating units. For each experimental iteration two supporting companies and a platoon opposing force will be identified. Identification of supporting units will be coordinated through the research sponsor. Because of the level of control that will be required over experimental conditions, a considerable amount of intervention will be necessary relative to pretest and posttest scenario design and scheduling for Dunn-Kempf exercises. Therefore, it is critical that unit commitments, in terms of personnel availability and scheduled exercises, be contracted in advance.

It is preferred that both participating companies come from the same battalion for two reasons. One reason is that both companies are likely to be at the major training areas at about the same time. While the time of pretest and posttest for both units need not correspond exactly, it is important that approximately the same interval exist between the two field

Figure 2. Procedural Approach to a Duplication of REALTRAIN
Transfer of Training Experiment



evaluations. It is assumed that the pretest and posttest will occur approximately three months apart due to conformity to normal unit training schedules; this should have the advantage of minimizing the learning effect of pretest on posttest, but bears the disadvantage of allowing more intervening "history" to occur between evaluations which might also account for observed differentials in performance. The second reason is that previous exposure to REALTRAIN and Dunn-Kempf should be roughly equivalent for the units and this will be more likely for two units in the same battalion.

In addition to enlisting the cooperation of units to participate in the experiment, it will be necessary to make arrangements to trace the training history (other than Dunn-Kempf training) for both units during the period of the pretest and posttest in order to identify alternative hypotheses to explain any changes found between initial and subsequent performance in REALTRAIN.

2. Train REALTRAIN evaluators. The major basis of comparison between experimental and control units will consist of tactical performance ratings which will be assigned by military evaluators. Military personnel who are selected as evaluators must be qualified to serve in the player positions they will be evaluating, namely company commander, platoon leader, and FIST chief. (The number of personnel and position requirements for different exercises are summarized in Table 1 at the end of the research plan). Prior to the REALTRAIN exercises, the assigned evaluators will go through a standardized training program which will acquaint them with the objectives of the evaluations in terms of the specific types of observations and evaluations that will be made. They will not, however, be told that they are involved in a transfer of training experiment.

The training will require from five to seven hours and should be conducted on the day before REALTRAIN exercises. Approximately 1/3 of the time will be used to review the REALTRAIN concept and procedures and to familiarize the evaluators with the terrain over which the exercises will be conducted. The remaining time will be devoted to reviewing the content of individual instrument items, identifying prerequisite conditions or cues for observation, and discussing data collection strategies for each measure. Contingencies relative to evaluations (e.g., equipment failure, or exercise anomalies) will be discussed along with various exercise characteristics which should be noted. Evaluator training will be conducted before each set of REALTRAIN exercises.

3. Collect reliability data. For the purposes of determining the reliability of the performance measures, two evaluators will be assigned to rate each player position over a series of ten to twelve REALTRAIN exercises (or until every measure has been rated at least eight times). This will be done for the positions of CO, the FIST chief and one platoon leader.

The evaluators will be instructed to apply the measures without collaboration. The series of paired ratings will provide data for the analysis of inter-rater reliability and individual measure stability. For the purposes of this investigation, it will be possible to use exercises involving different units and different training areas. The same data collectors should be used throughout.

4. Analyze data for reliability. Working with the paired ratings of evaluator teams, inter-rater reliability of measurement (for one exercise, all measures) will be determined using the Kuder-Richardson estimate of reliability,⁷ or other comparable techniques. This will be computed separately for each exercise and compared across exercises to determine the range and average reliability for the instrument. In addition, the stability of individual measures across exercises will be of concern in terms of screening measures for utility. Therefore, the two ratings on each measure will be correlated across the series of exercises (all exercises, one measure at a time) to determine the reliability, or stability, of the measure under varying tactical exercise conditions. To be considered adequate to the purposes of the experiment, reliability and stability indices should show correlations of at least .80. If the instrument does not meet the criterion for reliability, revision of the instrument should be undertaken with primary focus placed on measures showing the lowest stability. Then the revised instrument will need to be retested for reliability before proceeding with the transfer of training experiment.

5. Design REALTRAIN exercises. The research team will work with battalion training personnel to design a standardized problem (exercise) for both experimental and control units. This will entail developing generalized scenarios which are tailored to the available training area(s), the configuration of the opposing force (OPFOR) supporting the exercises, and assets of the units involved. The main design features of the exercise will include the locations of the offensive assembly area and objective, the general defensive deployment and strategy, the content of the battalion operations order (OPORD) to the company, the amount and types of indirect fire support that will be made available, and the manner of representing the Bn, CO, fire support coordinator (FSCoord), and other elements with which the company would be in communication. Exercises will be designed so that they can be conducted both in the field and on the Dunn-Kempf board game.

6. Conduct REALTRAIN pretests. The field pretest will consist of one REALTRAIN exercise for each of the experimental and control groups. During both exercises, incumbents of the five Dunn-Kempf player positions

⁷ J. P. Guilford, Fundamental Statistics in Psychology and Education, McGraw-Hill Book Company, 1965.

(company commander, three platoon leaders, and the FIST chief) will be evaluated using the performance evaluation instrument (Appendix A). At the same time objective data pertaining to the number and types of casualties and characteristics of the exercise (e.g., weather/visibility) will be collected by research personnel using the exercise record form (Appendix B).

7. Orient Dunn-Kempf trainers/controllers; design exercises. Dunn-Kempf training (the experimental treatment) will be conducted by military personnel familiar with administration of the game. (The assistance of a mobile training team of Dunn-Kempf advisors within division assets will be enlisted if this type of support is available. If not, personnel designated by the Bn/Bde will be trained in the administration of Dunn-Kempf). The Dunn-Kempf trainers will first be oriented to the requirements of the project and then will participate in the planning of the Dunn-Kempf exercises so that they approximately correspond to parameters established for the REALTRAIN exercises. In particular, the unit task organization, the OPFOR configuration, the combat support available, and the representation of adjacent and higher echelon elements should be comparable for both types of exercises. Because of the play of close air support (CAS) and air defense artillery (ADA) are not normally feasible in REALTRAIN, they will be excluded from Dunn-Kempf during the experimental treatment. By maximizing the correspondence between Dunn-Kempf and REALTRAIN exercises, it is assumed that the potential for transfer is optimized.

8. Train Dunn-Kempf evaluators. Three military personnel will be required to evaluate Dunn-Kempf exercises. One will evaluate the CO (who will be remote from the game board and control his organization based on maps, plans and communications), one data collector will evaluate FIST operations, and one will evaluate each of the three platoon leaders (in view of the protracted time over which play occurs, this latter requirement is considered feasible). As with REALTRAIN evaluators, Dunn-Kempf evaluators must be qualified for the positions they are evaluating by having had military service experience in this position.

The rules and play of Dunn-Kempf will be explained to the data collectors who will evaluate player performance in the Dunn-Kempf exercises. The major emphasis of the orientation will be training in the application of the tactical performance measures to Dunn-Kempf and interpretation of the performance standards in the simulated environment.

9. Conduct Dunn-Kempf training exercises. Prior to the REALTRAIN post-test, the experimental group will undergo a series of Dunn-Kempf training exercises which will be conducted according to procedures used in the preceding Dunn-Kempf study.⁸ After each exercise the game controller and the OPFOR player/controller will critique the exercise. This approximates the way that formal Dunn-Kempf training would normally be conducted within units. The three Dunn-Kempf evaluators will rate the five Dunn-Kempf players on each exercise and casualty data will be recorded by a research team

⁸ E. E. Miller and J. M. Bachta, op. cit.

observer. This will enable the tracing of performance gains over the exercises. However, none of the data collected during the exercises will be channeled back to the players. The number of training exercises will depend on the time and resources available however, a minimum of three should be used.

10. Conduct REALTRAIN posttests. Following Dunn-Kempf training, REALTRAIN exercises will be conducted for the experimental and control companies and performance of the CO, platoon leaders and FIST chief will be evaluated. As in the pretest, the exercises will be conducted on the same terrain and records of casualty assessments will be maintained. Also, the same evaluators should be used for both experimental and control units to increase the reliability of ratings for those exercises. As stated earlier, the evaluators should be "blind" to the designation of experimental and control units.

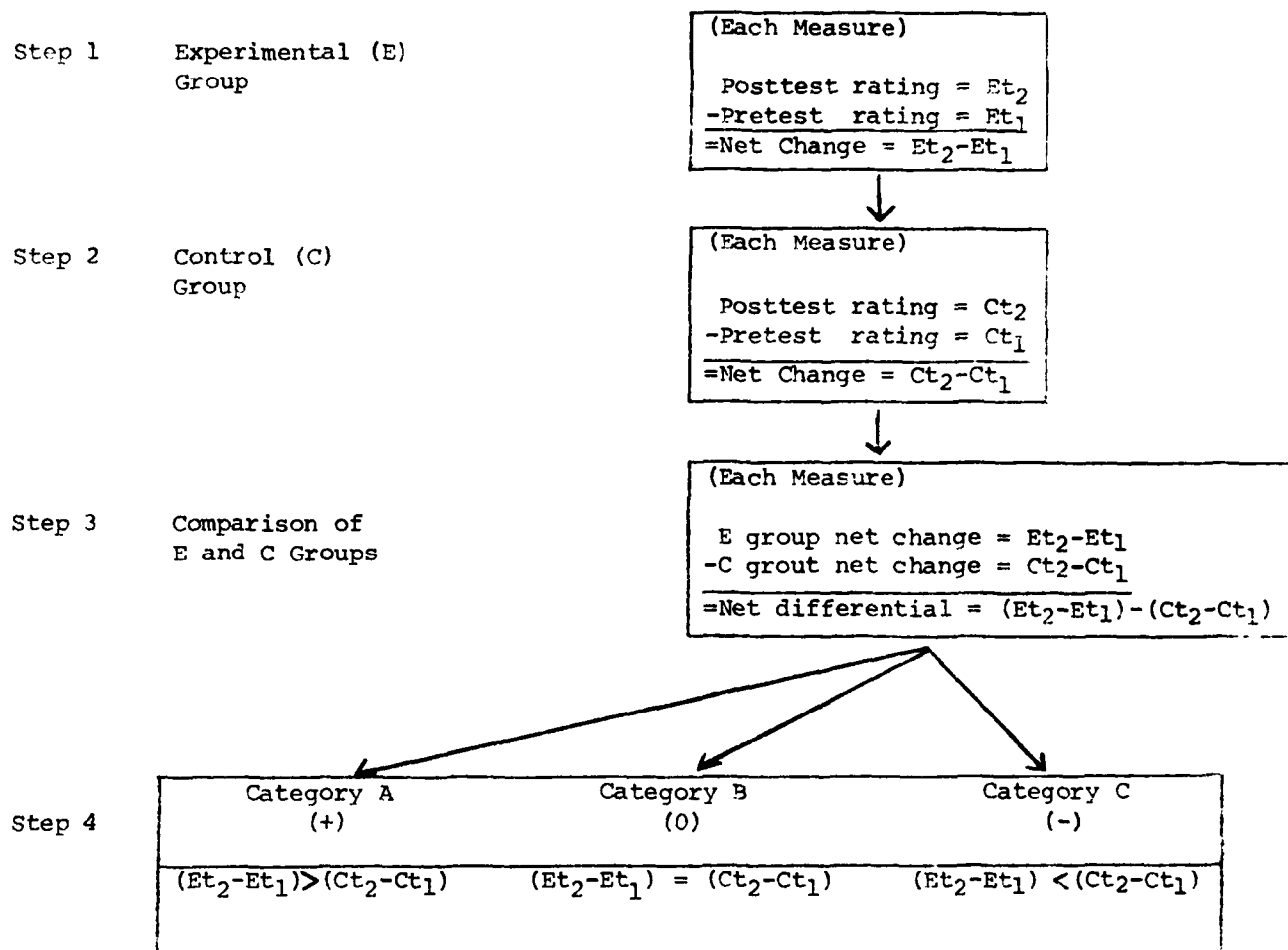
11. Analyze data for transfer of training. Because of the limited sample size anticipated and because of the psychometric problems involved in making quantitative comparisons of rating scales, data analysis will be largely restricted to the use of non-parametric statistics. Of foremost interest in the analysis will be the comparison of performance ratings between pre- and posttests in the criterion environment (REALTRAIN) for the experimental (E) and control (C) groups. The principal analysis that will be performed is depicted in Figure 3. Step 1 will consist of determining the net pretest-posttest change on each measure for the experimental group by subtracting the pretest rating on a measure from the posttest rating on the same measure. Step 2 will consist of determining the net change on each measure for the control group in the same manner. In Step 3, a net differential on each measure between E and C groups will be computed by subtracting the net change for the control group on a measure (Step 2) from the net change for the experimental group on the same measure (Step 1). Step 4 will consist of assigning each measure into one of three categories based on the sign and amount of net differential computed. The three categories, designated as A, B, and C, are defined as follows:

- (A) measures on which the experimental group showed a positive net change relative to the control group;
- (B) measures on which the net performance differential between experimental and control groups was zero;
- (C) measures on which the experimental group showed a negative net change relative to the control group.

Once all the measures have been assigned to a category, a Chi Square test will be computed to determine if the number of measures falling into each category differs significantly from what would be expected based on chance (Category A, $p=4/9$; Category B, $p=1/9$; Category C, $p=4/9$).⁹ If there is a significantly larger number of measures in Category A than in either Categories B or C, this would be an indication of a positive transfer effect for the experimental group.

⁹ S. Siegel, Nonparametric Statistics for the Behavioral Sciences, Mc-Graw Hill Book Company, 1956

Figure 3. Item Analysis of Performance Differentials
Between the Experimental and Control Groups



Where:

Et_1 = Pretest rating on a measure from the experimental group
 Et_2 = Posttest rating on a measure from the experimental group
 Ct_1 = Pretest rating on a measure from the control group
 Ct_2 = Posttest rating on a measure from the control group

Following this statistical test, the categories will be analyzed for content commonality among measures to determine the general tactical skill areas (e.g., movement techniques, use of IF) in which Dunn-Kempf training appeared to facilitate experimental group performance, and areas where there was no apparent advantage of Dunn-Kempf training. Actual ratings on the pretest for E and C groups will also be reviewed and compared to determine if there were possible "ceiling effects" in either group (initial performance so high that there was little room for improvement) that could have affected the net differentials. The category content analysis will be repeated for each iteration of E and C groups. The culmination of this analysis will be to trace whether measures tend to stay in the same differential effect categories (positive, zero, negative) or establish no pattern among experimental iterations. If a measure remains within the same effect category through three or more iterations, this will be a good indication of a probable transfer effect for the skill area tapped by the measure.

Another analytical approach may be possible if a sufficient number of units have been involved in the TOT experiment (6E and 6C). For each measure, a two-way analysis of variance (ANOVA) could be computed as shown in Figure 4. A mixed factorial model would be appropriate because E and C are independent while pretest and posttest scores are repeated measures.¹⁰ An indication of transfer of training would show up in a group by criterion score interaction effect.

Another analysis of interest is whether gains measured in Dunn-Kempf exercises correspond to gains on the same measures in REALTRAIN. This analysis will be accomplished by computing the correlation of the net changes between the first and last Dunn-Kempf exercise with the net pretest-posttest changes for the experimental groups across all measures.

A final analysis will compare the results of the objective outcome data (casualties assessed in each exercise) for the experimental and control groups. For the purposes of comparisons using outcome data, casualties assessed during each exercise will be converted to a weighted combat index (WCI). This is a technique developed and used in previous research efforts with REALTRAIN and Dunn-Kempf, and is used to derive a single outcome score for each side (offense and defense) which reflects the relative combat value of weapon systems destroyed. The formula for weighting the combat elements is adapted directly from the preceding Dunn-Kempf Study.¹¹

¹⁰

J.P. Guilford, op. cit.

¹¹

E. E. Miller and J. M. Bachta, op. cit.

Figure 4. Two-way ANOVA Model

CRITERION SCORE		
	Pretest	Posttest
Experimental Group	EG ₂ t ₁ , EG ₂ t ₁ ,	EG ₁ t ₂ , EG ₂ ,t ₂ ,
	EG ₃ t ₁ , etc.	EG ₃ t ₂ , etc.
Control Group	CG ₁ ,t ₁ , CG ₂ t ₁ ,	CG ₁ t ₂ , CG ₂ t ₂ ,
	CG ₃ t ₁ , etc.	CG ₃ t ₂ , etc.

Where: EG_n = rating for the nth experimental group on the particular measure

CG_n = rating for the nth control group on the measure

t₁ = pretest rating on the measure

t₂ = posttest rating on the measure

WCI' = 35 (# of tanks killed) + 25 (# of TOWs killed)

+15 (# of APCs killed) + 10 (# of Dragons killed)

+10 (# of squads killed) + 1 (# of infantry killed)

A comparative outcome index for the engagement is derived in the form of a WCI' ratio which is defined for this study as the ratio of losses inflicted to losses sustained:

$$\text{WCI' ratio} = \frac{\text{WCI' (OPFOR)}}{\text{WCI' (evaluated group)}}$$

For purposes of interpreting this ratio, the higher the ratio that is obtained, the better the outcome is for the evaluated group.

After a WCI' ratio is computed for each exercise, the pretest ratio for the experimental group will be subtracted from the pretest ratio for the same group to determine the net change in WCI' ratio for the experimental group. The same procedure will be followed to derive a net change for the control group. Comparison of the net pretest-posttest changes between groups will show whether the experimental group showed a positive, equivalent, or negative net differential relative to the control group. This analysis should be repeated for each iteration of E and C groups and results traced across iterations.

12. Prepare final report. A final report will be prepared which documents the procedures and results of the instrument reliability investigation and the transfer of training experiment. The report will include findings with regard to reliability of measurement and their implications for the interpretation of transfer of training. It will also discuss evidence for and against the inference of transfer and explain rationale used in interpreting results. Methodological insights gained relative to ways to improve control or sensitivity in the measurement of transfer will also be presented along with recommendations for further developmental work in this area.

TABLE 1

Research Plan
Support Requirements

RELIABILITY STUDY

Exercises:

(12) Company level combined arms REALTRAIN exercises opposed by combined arms augmented platoon.

Evaluator Personnel: (8-10)

(2) Company Commander evaluators
(4-6) Platoon Leader evaluators
(2) FIST Chief evaluators

(Evaluators should be currently serving in the military and have experience in the positions observed.)

Time Requirements:

7 days; one day for evaluator training; 6 days of exercises (2 per day).

Logistical Support:

(2) 1/4-ton trucks with dual net capability.
(8) back pack radios (PRC-77), field rations

TRANSFER OF TRAINING EXPERIMENT

Planning:

Battalion training personnel, 3-4 man-days.

Exercises:

Pretest: (2) Company level combined arms REALTRAIN exercises opposed by combined arms augmented platoon.

Experimental Treatment: (3 or more) Dunn-Kempf training exercises.

Posttest: (2) Company level combined arms REALTRAIN exercises opposed by combined arms augmented platoon.

TABLE 1 (Continued)

Research Plan
Support Requirements

Personnel:

REALTRAIN: (5)

- (1) Company Commander evaluator
- (3) Platoon Leader evaluators
- (1) FIST Chief evaluator

Dunn-Kempf: (5)

- (1) Controller
- (1) Assistant Controller
- (1) Company Commander evaluator
- (1) Platoon Leader evaluator
- (1) FIST Chief evaluator

(For both types of exercises, evaluators should be currently serving in the military and have experience in the positions observed.)

Time Requirement:

REALTRAIN: 6 days; one day for evaluator training and two days of exercises both for the pretest and posttest periods.

Dunn-Kempf: 4 days (minimum); one day for controller and evaluator training; one day each for three or more exercises.

Logistical Support:

REALTRAIN: (2) 1/4 ton trucks with dual net capability.
(3) back pack radios (PRC-77).

Dunn-Kempf: (1) Dunn-Kempf game board and kit.
(3) back pack radios or comparable wire net for five positions.

APPENDIX A

Dunn-Kempf/REALTRAIN

Performance Evaluation Form

T2/DCI-1

PERFORMANCE EVALUATION FORM (PEF)

Exercise # _____

Data Collector Name: _____

Date: _____

am _____ pm _____

Company Observed: _____

Element Observed: _____

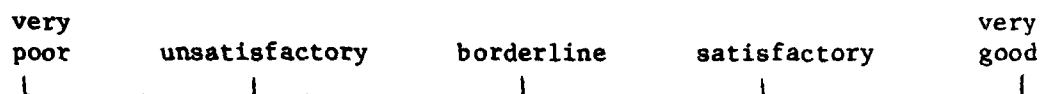
Position(s) Observed: _____

GENERAL INSTRUCTIONS

THE PURPOSE OF THIS DATA PACKET IS FOR USE IN ASSESSING THE TACTICAL PERFORMANCE OF A COMPANY TEAM IN A MOVEMENT TO CONTACT/ATTACK SITUATION. THE PRIMARY FOCUS FOR EVALUATION IS ON THE UNIT LEADERS (CO, PLT. LEADERS, FIST/FOs) AND HOW THEY EXERCISE THEIR COMMAND AND CONTROL FUNCTIONS. WE ARE INTERESTED IN WHAT THE UNIT LEADERS DO, HOW WELL THEY DO IT, OR WHAT THEY FAIL TO DO. INDIVIDUAL INFANTRYMEN OR VEHICLE CREW MEMBERS ARE NOT BEING DIRECTLY EVALUATED. FOR EXAMPLE, IF A UNIT LEADER WERE TO ASSIGN TARGETS FOR HIS UNIT, WE ARE INTERESTED IN EVALUATING THIS CONTROL FUNCTION BY ITSELF RATHER THAN EVALUATING WHETHER INDIVIDUALS ACTUALLY HIT THE TARGETS THEY ARE SHOOTING AT.

EACH ONE OF THE ITEMS (TACTICAL PERFORMANCE MEASURES) IN YOUR DATA COLLECTION PACKAGE WILL USUALLY INVOLVE THE FOLLOWING:

- INFORMATIONAL QUESTIONS that relate to the performance being observed. These questions should be answered/checked during the course of the tactical exercise. The questions usually pertain to factors or aspects of tactical performance that are emphasized in Army doctrine.
- OBSERVATION CONDITIONS are brief statements of when to observe or how to observe the performance being evaluated.
- PERFORMANCE RATING SCALE - toward the end of each exercise you will rate the overall performance of the individual or section that you are observing on each of the performance measures contained in the data collector package. For all measures, information recorded during an exercise should be reviewed and considered when rating overall performance on each measure. Each of the ratings will be made using the five point scale which follows.



The points on the scale are generally defined as follows: (Note that not all of the points in a definition will always apply to a measure.)

"BORDERLINE" - this is the balance point on the scale. It means that the overall performance on the measure could not consistently be judged satisfactory or unsatisfactory. The positive and negative behaviors cancelled each other out.

"SATISFACTORY" - means that most of the time the behavior that was appropriate was performed, the performance met accepted standards of tactical doctrine; it was in compliance with orders/directives; it was appropriate to the mission, terrain, and enemy situation.

"VERY GOOD" - means that almost all of the time, the requirements for a "satisfactory" rating are met, and in addition, the timing and execution of performance are clearly superior.

"UNSATISFACTORY" - means that most of the time the behavior that was appropriate was not performed; or if performed, it did not meet accepted standards of tactical doctrine; or performance was not in compliance with orders/directives; or performance was not appropriate to the mission, terrain, and enemy situation.

"VERY POOR" - means that almost all of the time the tactical performance was unsatisfactory in terms of the points mentioned above and/or showed some glaring deficiencies.

AT THE END OF MOST MEASUREMENT ITEMS THERE IS AN OPTION WHICH ALLOWS YOU TO INDICATE IF IT WAS NOT POSSIBLE TO OBSERVE ENOUGH TO RATE PERFORMANCE ON A MEASURE. IF THAT IS THE CASE, WE WOULD LIKE YOU TO IDENTIFY THE REASON THAT OBSERVATION OR MEASUREMENT WAS NOT POSSIBLE DURING THAT EXERCISE.

IN ALL INSTANCES WE WOULD ALSO LIKE YOU TO WRITE DOWN REMARKS ABOUT ANY CIRCUMSTANCES, WHICH YOU CONSIDER UNUSUAL OR EXCEPTIONAL, THAT AFFECTED PERFORMANCE.

Dunn-Kempf/REALTRAIN Measures

	<u>Position</u>
1. Operations Order	CO
2. Was the scheme of maneuver appropriate to the mission and terrain?	CO
3. Was fire plan developed which would support the scheme of maneuver?	CO/FIST
3.' Were direct fire weapon systems assigned priority of targets?	PL
4. Was the LD crossed on time?	CO
5. Was movement covered by overwatch, when appropriate?	CO/PL
6. Did the bounding element minimize exposure during movement?	PL
7a. Did overwatch positions selected maximize observation ability?	PL/FIST
7b. Did the overwatch positions selected enable fast and effective fire support for the maneuver elements?	PL/FIST
7c. Did the overwatch positions selected minimize vulnerability to enemy detection and fire?	PL/FIST
8a. Was CO kept informed of friendly situation?	CO
8b. Was CO kept informed of enemy situation?	CO
8'a. Was CO kept informed of friendly situation?	PL/FIST
8'b. Was CO kept informed of enemy situation?	PL/FIST
9. Was the FRAG Order complete and appropriate to the situation?	CO
10a. When engaged, did the engaged element take the appropriate action?	PL
10b. When engaged, did the supporting element take the appropriate action?	PL
11. Did weapon systems move after engaging enemy targets?	PL
12. After enemy contact is made, are elements controlled using appropriate fire and maneuver techniques?	CO/PL
13. Was suppressive fire used to support the final assault?	CO/PL

	<u>Position</u>
13.' Was suppressive fire used to support the final assault?	FIST
14. When contact was made, did platoon leader control the fire distribution of his elements?	PL
15. Did platoon leaders know the location of target reference points and preplanned fires?	PL
16. Was appropriate use made of target reference points and preplanned fires?	CO/PL/FIST
17. Was appropriate use made of indirect fire support?	CO/PL
17.' Was appropriate use made of indirect fire support?	FIST
18. Was communications security (COMSEC) maintained throughout the exercise?	CO/PL/FIST

1. Operations OrderEnemy Situation

_____ Did it include size, direction and type of enemy units?

Friendly Situation

Did it include:

_____ Battalion mission?

_____ Adjacent company missions?

Company Mission:

_____ Was mission stated?

Did it include: Who _____? What _____? Where _____? When _____? Why _____?

Execution

_____ Did the order include LD crossing time?

_____ Did the order identify the objective (N/A if no objective)

Were specific tasks assigned to:

_____ 1st Platoon?

_____ HAW Section?

_____ 2nd Platoon?

_____ GSR?

_____ 3rd Platoon?

_____ ENGR?

NOTE: Obtain phase lines, check points, preplanned fires, and target reference points from Commander's overlay.

Did scheme of maneuver include:

_____ Specific routes for each element?

_____ Phase lines?

_____ Check points?

_____ Overwatch assignments?

Fire Support Plan

Did it include:

_____ Preplanned fires?

_____ Target reference points?

_____ Priority of fires?

Command and Signal

Did it address:

____ CEOI?

____ Chain of command?

a) RATING: Operations Order

very poor	unsatisfactory	borderline	satisfactory	very good

2. Was the scheme of maneuver appropriate to the mission and terrain?

- . Did the CO specify general routes of advance for all elements? _____
- . Did the routes of advance selected by the CO for the elements take advantage of available cover and concealment? _____
- . Does maneuver scheme provide for continuous overwatch (when an overwatch movement technique is appropriate)? _____
- . Does the scheme of maneuver specify general location of overwatch positions? _____

a) **RATING:** Was the scheme of maneuver appropriate to the mission and terrain?

very poor	unsatisfactory	borderline	satisfactory	very good
--------------	----------------	------------	--------------	--------------

b) It was not possible to observe enough to make an evaluation: _____

Reason:

3. Was Fire Plan developed which would support the scheme of maneuver?

- . Were indirect fires planned to cover movement of the unit? _____
- . Did plan include use of smoke to mask elements moving through poorly covered areas? _____
- . Were preplanned fires designated on suspected enemy positions? _____
- . Were preplanned fires designated on, beyond, and on flanks of final objective (if appropriate)? _____

a) RATING: Was Fire Plan developed which would support the scheme of maneuver? (Consider also PL evaluator input regarding designation of target priorities for direct fire weapons).

very poor	unsatisfactory	borderline	satisfactory	very good

b) No fire plan was developed: _____

c) It was not possible to observe enough to make an evaluation: _____

Reason:

3'. Were direct fire weapon systems assigned priority of targets?

- a. Before the mission did the CO, PL or SOP prescribe target engagement priority for TOWs? Yes _____ No _____

If yes, indicate source: (CO, PL, SOP) _____:

target priorities

(1) _____

(2) _____

(3) _____

- b. Before the engagement were target priorities designated for tanks (main gun)? Yes _____ No _____

If yes, indicate source: (CO, PL, SOP) _____;

target priorities

(1) _____

(2) _____

(3) _____

- c. Before the engagement were target priorities designated for DRAGONS? Yes _____ No _____

If yes, indicate source: (CO, PL, SOP) _____;

target priorities

(1) _____

(2) _____

(3) _____

NOTE: This information should be sought from crews in the assembly area before the exercise. The information obtained should be submitted to the CO evaluator(s) during or immediately after the exercise.

4. Was the LD crossed on time?

- a) Time for crossing LD specified in OPORD: _____
- b) Was the LD crossed on time? Yes _____ No _____
- c) If no, what was the amount of deviation in minutes? _____
- d) It was not possible to observe the crossing of the LD and/or no time was specified: _____

Reason:

5. Was movement covered by overwatch, when appropriate?

(Oversight is normally appropriate when enemy contact is possible or expected.)

Observation #	Oversight was appropriate and used (✓)	Oversight was appropriate and not used (✓)	Remarks:
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			

a) RATING: Was movement covered by overwatch, when appropriate?

very poor unsatisfactory borderline satisfactory very good

b) It was not possible to observe enough to make an evaluation: _____

Reason:

6. Did the bounding element minimize exposure during movement?

Bound #	Movement technique utilized: (T,TO,BO)	Was technique appropriate to situation? (Y,N)	Was smoke used to mask bound through open areas? (Y,N, <u>Not Available</u>)	Was available terrain and foliage used to mask movement? (<u>Usually</u> , <u>Sometimes</u> , <u>Rarely</u>)	Did the element use rapid movement across open areas? (Y,N)	Did they maintain vehicle dispersion? (Y,N)	Did they bound to covered/concealed positions? (Y,N, <u>Not Available</u>)
1							
2							
3							
4							
5							
6							
7							
8							
9							
10							

a) RATING: (Element rated _____) Did the bounding element minimize exposure during movement?

very poor unsatisfactory borderline satisfactory very good

|-----|-----|-----|-----|

b) It was not possible to observe enough to make an evaluation: _____

Reason:

7. Execution of overwatch

Observation #	Did overwatch provide:			Did overwatch position minimize vulnerability to enemy fire (i.e., through advantageous use of available cover + concealment)? (Y,N)	Was overwatch mounted or dismounted? (M,D)	Considering preceding points, were better overwatch positions available? (Y,N)
	a) maximum observation of the route of the bounding element? (Y,N)	b) observation of most potential enemy positions to which the bounding element would be exposed (Y,N)	c) effective fields of fire and responsive fire support? (Y,N)			
1						
2						
3						
4						
5						
6						
7						
8						
9						
10						

a) RATING: Did overwatch positions selected maximize observation ability?

very poor unsatisfactory borderline satisfactory very good

|-----|-----|-----|-----|

- b) RATING: Did the overwatch positions selected enable fast and effective fire support for the maneuver element(s)?

very poor	unsatisfactory	borderline	satisfactory	very good

- c) RATING: Did the overwatch positions selected minimize vulnerability to enemy detection and fire?

very poor	unsatisfactory	borderline	satisfactory	very good

- d) It was not possible to observe enough to make an evaluation: _____

Reason:

8. Was CO kept informed of friendly/enemy situation?

Reportable situations: reaching check point(s)/phase line(s) - Friendly
 suffering crew casualties - Friendly
 experiencing equipment failure - Friendly
 initiating action - Friendly
 receiving fire - Friendly + Enemy
 encountering obstacle (e.g., minefield) - Friendly + Enemy
 sighting enemy vehicle(s)/position(s)/signature(s) - Enemy

	(a)	(b)	(c)	(d)	(e)	(f)
TRANS- MISSION #	Type of Report (E/F)	(✓) if initiated by CO	CO received sufficient information (Y,N)	If not sufficient did CO pursue better in- formation (✓ = Yes)	If routine reports were not timely did CO pursue information? (✓ = Yes)	Remarks:
1						
2						
3						
4						
5						
6						
7						
8						
9						
10						
11						
12						
13						
14						
15						
16						
17						
18						
19						
20						

a) RATING: Was CO kept informed of friendly situation?

very poor unsatisfactory borderline satisfactory very good

b) RATING: Was CO kept informed of enemy situation?

very poor	unsatisfactory	borderline	satisfactory	very good

c) It was not possible to observe enough to make an evaluation: _____

Reason:

8'. Was CO kept informed of friendly/enemy situation?

Reportable situations:

- reaching check point(s)/phase line(s) - Friendly
- suffering crew casualties - Friendly
- experiencing equipment failure - Friendly
- initiating action - Friendly
- receiving fire - Friendly + Enemy
- encountering obstacle (e.g., minefield) - Friendly + Enemy
- sighting enemy vehicle(s)/position(s)/signature(s) - Enemy

FIST {

- displacement of fire units (FA, mortars) - Friendly
- changes to priority of fire (Bn/TF and Bde) - Friendly

Type of communication:

- Commo present (e.g. - routine time checks)
- Control measure (check points, phase lines, objectives)
- Friendly Status - SITREP - (actions, locations, casualties, equipment failure)
- Enemy Action - SPOTREP - (vehicles, units, signature, receipt of fire, obstacles)
- Other

Observation #	Commo Type: (✓)					Did PL inform CO? (Y,N)	Was any of the information <u>inaccurate</u> ? (Y,N)
	Present	Control (friendly)	SITREP (friendly)	SPOTREP (enemy)	Other		
1							
2							
3							
4							
5							
6							
7							
8							
9							
10							
11							
12							
13							
14							
15							
16							
17							
18							
19							
20							

a) RATING: Was CO kept informed of friendly situation?

very poor	unsatisfactory	borderline	satisfactory	very good

b) RATING: Was CO kept informed of enemy situation?

very poor	unsatisfactory	borderline	satisfactory	very good

c) It was not possible to observe enough to make an evaluation: _____

Reason:

9. Was the FRAG Order complete and appropriate to the situation?

- . Did the circumstances of the engagement call for/require that a FRAG order be issued? Y _____ N _____ ? _____
- . Was a FRAG Order issued? Y _____ N _____
- . (If issued) did the FRAGO contain orders to all subunits _____, including fire support plan _____, and priorities of fires? _____
- . (If issued) did the FRAGO include control measures? _____
- . (If issued) did the FRAGO concentrate Company Team combat power, _____ and make maximum use of available direct and indirect fire? _____

a) RATING: Was the FRAG Order complete and appropriate to the situation?

very						very
poor	unsatisfactory	borderline	satisfactory			good

b) (If issued) It was not possible to observe enough to make an evaluation: _____

Reason: _____

PL

10. When engaged, did (a) engaged element, (b) supporting element take appropriate action?

Engagement #	Not observed or not applicable (✓)	Engaged by Direct Fire				Engaged by Indirect Fire			Action appropriate? (Y,N)
		return fire	deploy	report	request IF/DF fire	button up	move/seek cover	report	
1 Engaged Elem.									
1 Support Elem.									
2 Engaged Elem.									
2 Support Elem.									
3 Engaged Elem.									
3 Support Elem.									
4 Engaged Elem.									
4 Support Elem.									

a) RATING: Did element engaged take appropriate action?

very poor	unsatisfactory	borderline	satisfactory	very good

b) RATING: Did supporting element (if applicable and observed) take appropriate action?

Very poor	unsatisfactory	borderline	satisfactory	very good

c) It was not possible to observe enough to make an evaluation: _____

Reason:

11. Did weapon systems move after engaging enemy targets?

Observation #	Weapon System moved after shooting (Y,N)	Instances where it would have been inappropriate to move (✓)
1		
2		
3		
4		
5		
6		
7		
8		
9		
10		

a) RATING: Did weapon system(s) move after engaging enemy target(s)?

very					very					
poor	unsatisfactory	borderline	satisfactory		good					
<table border="1" style="width: 100%; height: 20px;"> <tr> <td style="width: 20%;"></td> <td style="width: 20%;"></td> <td style="width: 20%;"></td> <td style="width: 20%;"></td> <td style="width: 20%;"></td> </tr> </table>										

b) It was not possible to observe enough to make an evaluation: _____

Reason: _____

12. After enemy contact is made, are elements controlled using appropriate fire and maneuver techniques?

- . Did the unit (elements) that made contact establish a base of fire (BOF) to provide direct and indirect fire support for maneuver?

Y _____ N _____

Remarks: _____

- . Did the maneuver unit at any time mask the overwatching fires or advance beyond the effective supporting range of the BOF? Y _____ N _____

Remarks: _____

- . Was fire and maneuver continuously coordinated between the BOF and maneuver elements? (i.e., were the elements in constant radio contact)?

Y _____ N _____

Remarks: _____

- . Were enemy positions within reach suppressed or destroyed so that the maneuvering element(s) could continue to advance? Y _____ N _____

Remarks: _____

- a) **RATING:** After enemy contact is made, are elements controlled using appropriate fire and maneuver techniques?

very poor	unsatisfactory	borderline	satisfactory	very good

- b) It was not possible to observe enough to make an evaluation: _____

Reason: _____

11. Was suppressive fire used to support the final assault?

. Did the leader observed have assets available:

Non-organic direct fire?_____ Indirect fire?_____

. Was suppressive fire used on objective/enemy positions prior to and during assault:

Non-organic direct fire?_____ Indirect fire?_____

. Was suppressive fire shifted to the rear and flanks of the enemy during the assault?

Yes_____ ; No_____ ; Wasn't Appropriate_____ ; Could not determine_____.

a) RATING: Was suppressive fire used to support the final assault (given the assets that were available)? Note: Incorporate input from FIST evaluator.

very poor	unsatisfactory	borderline	satisfactory	very good

b) It was not possible to observe enough to make an evaluation:_____

Reason:

FIST

13'. Was suppressive fire used to support the final assault?

. Were IF assets available to CO____, PL____ at the time of the assault?

. Did FIST chief utilize suppressive IF on objective/enemy positions in support of the assault? (Y,N) _____

HE? _____

Smoke? _____

a) RATING: Was suppressive IF used to support the final assault (given the assets that were available)?

very					very
poor	unsatisfactory	borderline	satisfactory		good

b) It was not possible to observe enough to make an evaluation: _____

Reason:

NOTE: Submit this information to CO evaluator immediately after the exercise for his overall evaluation of this item.

14. When contact was made, did platoon leader control the fire distribution of his elements?

. At time of contact, did PL know where his elements were? Y____ N____

. Did the PL assign specific targets or sectors to his squads or sections? Y____ N____

. Were most engagements "one-on-one" type of situations? Y____ N____

. Did platoon have SOP for fire distribution on contact? Y____ N____

a) RATING: When contact was made, did platoon leader control the fire distribution of his elements?

very poor	unsatisfactory	borderline	satisfactory	very good

b) It was not possible to observe enough to make an evaluation: _____

Reason:

15. Did platoon leaders know the location of target reference points and preplanned fires?

a) Yes _____ No _____

Conditions for observation:

- Platoon leader calls for TRP/PF fires.
- After exercise is completed ask platoon leader for TRP/PF locations, or look at PL's overlay.

b) It was not possible to observe enough to make an evaluation: _____

Reason:

16. Was appropriate use made of target reference points (TRP) and preplanned fires (PF)?

. Were TRPs selected/requested by CO? Y ☐ N ☐
 PL? Y ☐ N ☐

. Were TRPs located near distinguishable landmarks, such as:

road junctions? ☐
 lone tree or rock formation? ☐
 hill top? ☐
 tree line? ☐
 Other (specify)

. Was there generally one TRP per 1000 meters (i.e., each grid square)? ☐

. Were TRP/PFs visible from successive friendly positions? ☐

. Were TRPs used to control maneuver or fire distribution? ☐

. When TRP/PF was near a target, was it used? Y ☐ N ☐ (Keep a tally if there are multiple observations.)

a) **RATING:** Was appropriate use made of TRPs and PFs?

very					very
poor	unsatisfactory	borderline	satisfactory		good
☐ ☐ ☐ ☐ ☐					

b) It was not possible to observe enough to make an evaluation: ☐

Reason:

17. Was appropriate use made of indirect fire (IF) support?

- . Were good opportunities to use IF overlooked? _____ How many times? _____
- . Were IF missions used when not needed? _____ How many times? _____
- . Were appropriate means/assets requested for IF missions? _____
- . Was fire support coordinated efficiently? _____

Obtain the following from the FIST evaluator after the exercise:

- . Number of smoke missions available to the unit _____
- . Number of smoke missions called by the unit: _____
- . Number of HE missions available to the unit: _____
- . Number of ARTY missions called by the unit: _____
- . Number of mortar missions called by the unit: _____

- a) **RATING:** Was appropriate use made of indirect fire (considering the IF assets available and used)?

very				very
poor	unsatisfactory	borderline	satisfactory	good
----- ----- ----- -----				

- b) It was not possible to observe enough to make an evaluation: _____

Reason:

17. Was appropriate use made of indirect fire (IF) support?

- . How many smoke missions were available to the unit? _____
- . How many smoke missions were called by the unit? _____
- . How many HE missions were available to the unit? _____
- . How many ARTY missions were called by the unit? _____
- . How many mortar missions were called by the unit? _____

NOTE: The above information should be shared with the CO and PL evaluators immediately after the exercise.

- . Were good opportunities to use IF overlooked? _____ How many times? _____
 - . Were IF missions used when not needed? _____ How many times? _____
 - . Were appropriate means/assets requested for IF missions? _____
 - . Was fire support coordinated efficiently? _____
- a) **RATING:** Was appropriate use made of indirect fire (considering the IF assets available and used)?

very poor	unsatisfactory	borderline	satisfactory	very good

- b) It was not possible to observe enough to make an evaluation: _____

Reason:

18. Was communications security (COMSEC) maintained throughout the exercise?

. Were violations of COMSEC observed during the exercise (e.g., transmission of unit coordinates or names of individuals or organizations in clear)?

Y _____ N _____

. If observed, how many COMSEC violations occurred and of what type?

Number observed: _____

Type(s): _____

a) **RATING:** Was communications security (COMSEC) maintained throughout the exercise?

very poor	unsatisfactory	borderline	satisfactory	very good

b) It was not possible to observe enough to make an evaluation: _____

Reason:

APPENDIX B

Exercise Record Form

Exercise Record Form (ERF)

Exercise # _____

Date _____ am _____ pm _____

Offense: (Friendly)

Unit designation: _____

Mission: _____

Unit configuration/assets: _____

General method of execution/deployment: (e.g. - column, flanking, etc.)

Defense: (OPFOR)

Unit designation: _____

Mission: _____

Unit configuration/assets: _____

General method of execution/deployment: (e.g. - in depth, dispersed, clustered, etc.) _____

Force ratio: _____

Characterization of terrain utilized: _____

Weather conditions and visibility: _____

Notations regarding control systems: (e.g. - deviations such as equip-
ment problems or personnel shortages, changes in procedures, etc.) _____

Outcome data:

Offensive tank casualties: _____

Offensive APC casualties: _____

Offensive TOW casualties: _____

Offensive dismounted infantry casualties: _____

Offensive other (specify) weapon system casualties: _____

Indirect Fire Data:

Offense:

Number of missions called: _____

Number of rounds delivered: _____

Number of instances of suppression: _____ (total)

- mortar _____

- arty _____

Number of Arty casualties achieved by target type:

tank: _____

APC: _____

personnel: _____

other (specify): _____

Defense:

Number of missions called: _____

Number of rounds delivered: _____

Number of instances of suppression: _____

- mortar _____

- arty _____

Number of Arty casualties achieved by target type:

tank: _____

APC: _____

personnel: _____

other (specify): _____

Note: Get roster of vehicle types and RT numbers for each side from NCS.