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

Research Report 1679

An Initial Evaluation of a Simulation-Based Training Program for Army National Guard Units

Theodore M. Shlechter and David W. Bessemer
U.S. Army Research Institute

Paul Nesselroade, Jr., and James Anthony
University of Louisville

June 1995

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

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

EDGAR M. JOHNSON
Director

Technical review by

Lieutenant Colonel William C. Dorman
Robert A. Wisher

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	1995, June	Final Jan 94 - Jan 95

4. TITLE AND SUBTITLE
An Initial Evaluation of a Simulation-Based Training Program for Army National Guard Units

5. FUNDING NUMBERS
63007A
793
2124
H01

6. AUTHOR(S)
Shlechter, Theodore M.; Bessemer, David W. (ARI); Nesselroade, Jr., K. Paul; Anthony, James (University of Louisville)

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
U.S. Army Research Institute for the Behavioral and Social Sciences
Armored Forces Research Unit
ATTN: PERI-IK
Fort Knox, KY 40121-5620

8. PERFORMING ORGANIZATION REPORT NUMBER
ARI Research Report 1679

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Research Institute for the Behavioral and Social Sciences
5001 Eisenhower Avenue
Alexandria, VA 22333-5600

10. SPONSORING / MONITORING AGENCY REPORT NUMBER
--

11. SUPPLEMENTARY NOTES
K. Paul Nesselroade, Jr., and James Anthony are graduate students at the University of Louisville working as consortium fellows at the U.S. Army Research Institute's Armored Forces Research Unit.

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

12b. DISTRIBUTION CODE
--

13. ABSTRACT *(Maximum 200 words)*

This research effort was designed to provide initial empirical information needed to examine the Reserve Component Virtual Training Program's (RCVTP's) instructional effectiveness. A multimethod-multisource research strategy was used to address this objective. Observers collected data from nine units, who executed 45 tables (exercises); fourteen RCVTP instructors completed standard rating forms regarding the performance of 38 armored force units; and 280 training participants completed Likert-scale items regarding their training experience. Data from the different methods indicated that the units further developed their collective tactical skills across the training period. They took significantly less time, made fewer errors, and needed less coaching as their training progressed. The instructors indicated that most units had a greater likelihood of becoming more proficient in critical subtasks than either not improving or becoming less proficient. The participants claimed that they were more proficient after training than before. The RCVTP should continue to be used to train Army National Guard armored units.

14. SUBJECT TERMS
Armor training
Simulation networking
Simulation-based training

15. NUMBER OF PAGES
81

16. PRICE CODE
--

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

Research Report 1679

An Initial Evaluation of a Simulation-Based Training Program for Army National Guard Units

Theodore M. Shlechter and David W. Bessemer
U.S. Army Research Institute
Paul Nesselroade, Jr., and James Anthony
University of Louisville

Armed Forces Research Unit
Barbara A. Black, Chief

Training Systems Research Division
Jack H. Hiller, Director

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

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

June 1995

Army Project Number
2O236007A793

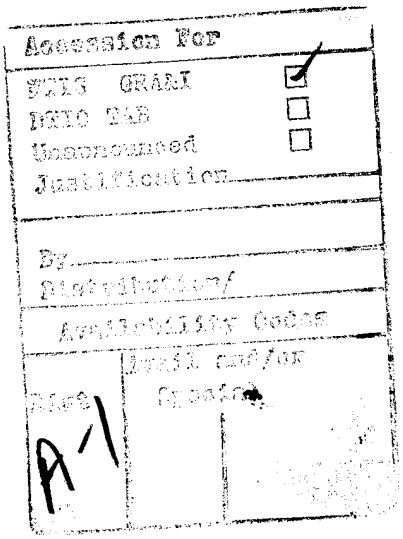
Training Systems and Education

FOREWORD

Army National Guard (ARNG) units have become an increasingly important element of post cold-war combat power. ARNG soldiers must be trained for their new roles in the post cold-war military. To support the needed training for armor ARNG training, Congress has provided funding to establish a Reserve Component Virtual Training Program (RCVTP). This training program involves structured exercises conducted in a high technology simulation training environment--the SIMulation NETworking (SIMNET) training system--to provide ARNG armor units with intensive training experience during their weekend drills or annual training (AT) periods.

The U.S. Army Research Institute for the Behavioral and Social Sciences (ARI), the Advanced Research Projects Agency (ARPA), the National Guard Bureau (NGB), the U.S. Army Armor Center (USAARMC), and Fort Knox joined efforts (Memorandum of Agreement entitled "National Guard Armor Simulation Center," April 1993) to develop and implement the RCVTP. The ARI Armored Forces Research Unit at Fort Knox accomplished training research and development for the RCVTP through a contract effort entitled "Simulation-Based Multiechelon Training Program for Armor Units (SIMUTA)," as part of the Research Task entitled "Strategies for Training and Assessing Armor Commanders' Performance with Devices and Simulations (STRONGARM)."

This research describes an initial assessment of the RCVTP's training value. It integrates data from observers' reports, instructors' judgments of performance, and training participants' questionnaire responses. The information in this report has been provided to training developers and instructors in the 16th Cavalry Regiment at Fort Knox. It will also be useful to all personnel involved in the development and implementation of structured simulation-based training.



EDGAR M. JOHNSON
Director

AN INITIAL EVALUATION OF A SIMULATION-BASED TRAINING PROGRAM FOR ARMY NATIONAL GUARD UNITS

EXECUTIVE SUMMARY

Requirement:

Army National Guard (ARNG) units have become an increasingly important element of post cold-war combat power. ARNG soldiers must be trained for their new roles in the post cold-war military. To support the needed ARNG training, Congress has provided funding for establishing a Reserve Component Virtual Training Program (RCVTP) that uses the available training technologies at Fort Knox, KY, including the Simulation Networking (SIMNET) system. This research effort has been designed to provide initial empirical information needed to examine the RCVTP's instructional effectiveness.

Procedure:

This research involved using trained observers, RCVTP instructors (observer/controllers--O/Cs), and RCVTP participants to assess the RCVTP's instructional effectiveness during the initial developmental trials of the program's SIMNET training exercises. The exercises (called tables) provide structured training for tactical fundamentals, offensive missions, and defensive missions. The observers collected data from nine units, who executed 45 tables. Seven units executed four or more tables. The observers recorded the: (a) time the units took to complete a table; (b) tactical performance of the units; and (c) coaching provided by the O/Cs. They also rated units' table performance and recorded information regarding the after-action reviews (AARs) that followed table execution.

Fourteen RCVTP instructors completed standard rating forms. They identified those subtasks in which units were proficient and those in which units needed to improve. Data for these ratings were collected from 38 armored force units. These units executed 187 tables, comprising 32 fundamental, 89 offensive, and 66 defensive tables. More than 75% of the units completed 4-6 tables in a two-day training period using SIMNET.

Two hundred eighty training participants completed Likert-scale items regarding their training experience. These soldiers estimated their proficiency before and after training on a 1-7

scale, with 7 being extremely proficient and 1 not at all proficient. The participants also gave their opinions regarding different aspects of this training program.

Findings:

Data from the different methods indicated that the units further developed their collective tactical skills across the training period. They took significantly less time, made fewer errors, and needed less coaching as their training progressed. The instructors indicated that most units had a greater likelihood of getting more proficient in critical subtasks than either not improving or getting worse in them. The participants claimed that they were more proficient after training than they were before training. These last two findings were more pronounced for the platoons, which were comprised of the highest proportion of ARNG units. Also, the leaders from the ARNG armor companies felt, after training, as competent as did leaders from active component units.

These improvement trends occurred despite the fact that the successive tables became increasingly more difficult. RCVTP was thus shown to be an effective instructional program for ARNG training proposes.

Utilization of Findings:

This report has ramifications for military trainers, evaluators, and instructional designers. Empirical support has been provided for using the RCVTP to train ARNG units. Also, this report has further delineated the advantages of using a variety of methodologies to conduct naturalistic evaluations of high technology training systems.

AN INITIAL EVALUATION OF A SIMULATION-BASED TRAINING PROGRAM FOR ARMY NATIONAL GUARD UNITS

CONTENTS

CONTENTS (Continued)

LIST OF TABLES

Page

LIST OF FIGURES

AN INITIAL EVALUATION OF A SIMULATION-BASED TRAINING PROGRAM FOR ARMY NATIONAL GUARD UNITS

Introduction

This report describes an initial evaluation of the Reserve Component Virtual Training Program (RCVTP). This training program involves structured exercises conducted in a high technology simulation training environment--the SIMulation NETworking (SIMNET) training system--to provide Army National Guard (ARNG) armor units with intensive training experience during their weekend drills or Annual Training (AT) periods.

The instructional design and development processes associated with the RCVTP were accomplished by the Simulation-based Multi-echelon Training Program for Armor Units (SIMUTA) contractor team, which consisted of a consortium of military subject-matter experts, instructional designers, and evaluators. This contractual effort was monitored by the U.S. Army Research Institute for the Behavioral and Social Sciences (see Hoffman, Graves, & Koger, 1994 for a detailed description of the SIMUTA project).

The RCVTP

Need for the RCVTP. ARNG units have become an increasingly important element of post cold-war combat power. These units, however, have limited training resources and time, with only 39 days allocated for training per year, including just 15 days for AT. Congress has thus provided funding for establishing the RCVTP at Fort Knox, KY.

The RCVTP's instructional framework. This program's primary goal involved having ARNG units experience National Training Center (NTC)-like missions in a time-compressed manner. Hence, the courseware associated with the RCVTP was based on two cornerstone battalion missions (movement to contact and defense in sector) frequently used at the NTC.

Providing NTC-like training in a time-compressed manner involved utilizing the available high technology training simulation systems at Fort Knox, KY. As indicated, the primary simulation utilized by the RCVTP was the SIMNET system. The Janus constructive simulation system was also used to support battalion staff training exercises. Janus exercises, however, are not discussed in this report, as this evaluation only dealt with the SIMNET portion of the RCVTP.

SIMNET is a real-time distributed interactive battle simulation that provides collective tactical training for units from platoon to battalion levels (see Garvey & Radgowski, 1988; Shlechter, Bessemer, Rowatt, & Nesselroade, 1994; Turcek, Campbell, Myers, & Garth, 1994 for a further description of this system). A major problem with SIMNET is that it has not been supported with a structured instructional program. Bessemer

(1991) has thus suggested that this system's effectiveness is related to the training techniques employed by its instructors. Some armor units might then benefit from SIMNET while others might not. The SIMUTA instructional design team thus had to develop a structured set of SIMNET training exercises (tables) in order to maximize SIMNET's capabilities to train ARNG units.

C. H. Campbell, R. C. Campbell, Sanders, & Flynn (1994) have noted that structured simulation-based training is characterized by focusing on specific training objectives. All training tables were thus structured to have units perform actions (critical subtasks) associated with specific training objectives and cues. Examples of critical subtasks included: (a) reaching the starting point on time; (b) executing fires when the enemy crosses the trigger line; and (c) conducting displacement as directed.

Attempting to maximize SIMNET's capabilities also entailed creating "turn-key" sets of training support materials that allowed units to focus on maneuver execution. That is, operation orders, overlays and other planning materials were provided to units in advance of their RCVTP training. Correspondingly then, units would focus on executing the training tasks rather than on managing the training exercises (C. H. Campbell et al, 1994). Also, units would spend more time at the SIMNET facility on executing missions than they would on planning and preparing for the different missions.

The RCVTP was thus comprised of structured tables for SIMNET-based training of collective tasks. Approximately one hundred training tables were created for this training program. This program also included a two-hour familiarization course designed for crews experiencing SIMNET for the first time (Turcek, C. H. Campbell, Myers, & Garth, 1994) The platoon and company training tables, which are the focus of this evaluation, are discussed in the next section.

Structure of the platoon and company training tables. The SIMUTA team created 18 armor platoon training tables, 18 mechanized infantry (mech) platoon training tables, 12 scout platoon training tables, 18 armor company and 18 company team training tables. (Additional training tables were created for cavalry and battalion-level exercises.) Based on Army training recommendations (Morrison & Holding, 1990), these training tables were structured to flow from the cornerstone battalion missions in a "crawl-walk-run" sequence of learning. This sequence meant that the lessons were designed to become more intricate as the training progressed. Three crawl (fundamental) training tables and 15 walk-run (9 offensive and 6 defensive) training tables were developed for the armor company, company team, armor platoon, and mech units. For the scout platoon, 3 crawl and 9 walk-run (6 offensive and 3 defensive) training tables were developed.

These training tables were structured so that units faced

more difficult critical subtasks as they progressed from the fundamental training tables to the more complicated offensive and defensive training tables. Some easy critical subtasks were also included in the latter training tables so that the participants would have the opportunity to repeat training on selected tasks (C. H. Campbell et al., 1994). The fundamental training tables thus consisted of units engaging few if any enemies, moving across easily negotiable terrain, and following the orders as briefed; while the later training tables consisted of missions which included having units cross difficult terrain and respond to FRAGOs (fragmentary orders, which are used to change the mission). The platoon and company training tables were designed to become increasingly difficult as units progressed through them. The difficulty of the different critical subtasks was determined by the military subject-matter experts (SMEs) associated with the SIMUTA instructional design team (C. H., Campbell et al., 1994).

Each training table was also designed to be conducted in two hours. One half-hour was spent by units on preparing for the mission, 1 hour on executing the mission, and another 1/2 hour on participating in an after-action review (AAR) of the exercise. The times for the preparation and AARs might vary depending on the mission. The AAR process for a company, for example, might last 45 minutes with 20 minutes being spent on a platoon AAR and 25 minutes on a company AAR. The latter AARs would only include the company commander, executive officer, platoon leaders, and tank commanders.

Implementation of the RCVTP. The RCVTP is being implemented by a dedicated team of approximately 20 observer/controllers (O/Cs) at Fort Knox. These O/Cs, who are primarily Active Component military personnel ranging in rank from Sergeant First Class to Lieutenant Colonel, have the following responsibilities associated with implementing this program:

1. Preparing personnel for a rotation at the RCVTP by visiting them at their home station. During such visits, the O/Cs provide the unit with all required training support materials, including overlays and operation orders associated with completing RCVTP exercises.

2. Providing units with a preview of the training table at the O/C station (see next paragraph for a description of this station). This preview, which occurs during the time for the preparation phase, involves giving the unit an operation order, providing them with a quick overview of the battlefield terrain, and answering their questions about the mission.

3. Overseeing execution of the RCVTP training tables. This responsibility involves helping units in table preparation and monitoring their table execution. The O/C may, if needed, help units to complete the training table.

4. Facilitating the AARs. The O/Cs have been trained by the SIMUTA instructional design team to conduct these AARS by following a specific preplanned agenda and by facilitating discussion among the group. Hoffmann et al. (1994) have noted that this training should help the O/Cs to conduct quality AARs with maximum payoffs in productive time.

5. Completing the Take Home Package (THP). The THP consists of the O/Cs' observations regarding training performance. This package is sent to the units' home stations following their participation at the RCVTP, and is used to assess their future training needs (Turcek et al., 1994). The THP will be further discussed in the sections on the observers' assessment.

O/Cs can observe performance and lead the AARs from their stations at the SIMNET facility. These stations contain a plan view display (PVD), tactical radios, stealth vehicle controls, and audiovisual recording and replay system helping the O/Cs to perform these duties. The stealth vehicle, for example, provides a direct view of the battlefield from an invisible vehicle moving on or above the terrain.

RCVTP training schedule. These O/Cs have also been asked to comply with the following training schedule when implementing the RCVTP for a typical weekend drill. The schedule provides each ARNG unit with approximately 2 days of SIMNET training. The first two hours are spent in helping units to become, if needed, familiar with the SIMNET equipment, which will include completing the familiarization course. The next half-day is to be spent completing the fundamental training tables. When the O/Cs are confident that the units have mastered their fundamental subtasks, then they will complete the offensive or defensive training tables. It must be noted that the RCVTP set of training tables provides a training library or menu with a recommended sequence rather than a prescriptive training matrix. That is, the O/C and unit leaders have flexibility to choose tables that meet their training purposes within the available time.

Objective of This Evaluation

This investigation has thus been designed as an initial evaluation of the RCVTP's effectiveness for training ARNG units. Effectiveness has been defined as helping these units to further develop their collective tactical skills.

Evaluation Approach

Need for a multimethod-multisource evaluation. This objective presents several interesting challenges, of which the primary challenge is to determine the most appropriate method(s) for collecting data. While SIMNET has the previously discussed observational aids, automatic performance measuring devices are not yet fully operational (Meliza, Bessemer, & Tan, 1994). Researchers must then collect data through observations (Gound &

Schwab, 1988), questionnaires (Brown, Pishel, & Southard, 1988), or instructor ratings (Bessemer, 1991; Shlechter, Bessemer, & Kolosh, 1991).

Each method is potentially problematic. Observational methods are labor intensive and limit the sample size. Bessemer (1991) has noted that problems (e.g., lack of statistical power) associated with small sample sizes are exacerbated by assessing collective training devices with intact military units used as the sampling units. Questionnaires can be tainted by the participants' inability to report accurately the effects of the training device on their performance. The accuracies of self-reports have been hotly debated by psychologists (see Burnside, 1982; Nisbett & Wilson, 1977; Herrmann, 1982). Instructor ratings may be contaminated by the expectations or biases held by the instructors (Cook & Campbell, 1979).

The previously cited instrumentation problems become more critical in naturalistic evaluations, which tend to be quasi-experimental or pre-experimental designs (Shute & Regian, 1993). A defining characteristic of such designs is that they lack rigorous experimental controls. Without such controls, it becomes difficult to remove the effects of extraneous variance from the data (Cook & Campbell, 1979).

Researchers, who prefer data to have high internal validity, are often compelled to conduct quasi-experimental evaluations of high technology training systems. The RCVTP training managers, for example, felt that conducting an evaluation with experimental controls would encroach upon their training program. That is, they wanted the training conditions for the RCVTP's formative evaluation to be very similar to the training conditions for the fully implemented program. The present evaluators were thus further challenged to conduct a quasi/pre-experimental evaluation of the training program.

Previous investigations of SIMNET's effectiveness have also involved the use of quasi/pre-experimental designs (e.g., Bessemer, 1991; R. E. Brown, et al., 1988; Shlechter, Bessemer, & Kolosh, 1991). Bessemer tried to reduce confounded effects by removing statistically the effects of extraneous variables from his data. This evaluation consisted of obtaining instructor ratings for 1705 Armor Officer Basic (AOB) students of which 1059 did not receive SIMNET training and 646 did. The former group of AOB students needed more practice on elementary contact exercises during field maneuvers than did their SIMNET trained counterparts. Multiple regression techniques were used to help remove the effects associated, for example, with instructor biases. As indicated, this evaluation involved a large sample size, which was not possible for the present evaluation. Also, statistical adjustments do not fully substitute for experimental controls.

A multimethod evaluation approach has been suggested as a

technique for circumventing the cited limitations with naturalistic evaluations(e.g., Denizen, 1978 as cited in Patton, 1987; Cook & Campbell, 1979). Areas of agreement between methods would boost confidence in the data's internal and construct validity (Cook & Campbell). Denizen has also suggested that naturalistic evaluations include a variety of investigators. Each investigator could provide a different perspective regarding the training situation. Independent observers may view a subject's performance differently from an instructor who is part of the training process. Judgments (assessments) of performance should then come from a variety of sources.

A "multimethod-multisource" evaluation strategy was adopted in order to address this evaluation's primary objective. This strategy involved conducting, simultaneously, different assessments of performance. The selection of methods and sources for these assessments was based on the previous SIMNET evaluations. These different assessments consisted of observations by evaluators, instructors' (RCVTP O/Cs') judgments of performance, and participants' questionnaire responses. These methods are further delineated in the sections dealing with Assessment A (observations by researchers), Assessment B (O/Cs' judgments) and Assessment C (questionnaire responses).

Need to sample numerous criterion measures. The multimethod-multisource strategy could also be useful for helping this evaluation to meet another challenge--sampling adequate criterion measures--associated with assessing high technology training systems. Shute and Regian (1993) have noted that sampling adequate criterion measures has been a problem which has plagued evaluations of high technology training systems, especially those systems designed to help students to become proficient in performing complex tasks. Hence, this evaluation was also designed to sample numerous criterion measures.

Participants to be sampled. Most of the participants for this evaluation were part of the developmental trials phase of the SIMUTA effort, which used the initial courseware. This phase took place during the Winter and Spring of 1994. The other phases of this contract included refinement trials in which revised courseware was utilized. All units volunteered to participate in this evaluation.

Assessment A: The Observers' Reports

This assessment was conducted by evaluators from the U.S. Army Research Institute for the Behavioral and Social Sciences (ARI). These evaluators were independent of both the instructional design and training process.

Method

Participants. Nine units were observed. Seven of these units--five ARNG and two active units--were observed during the

developmental trials. These units consisted of three armor companies, two armor company teams, and two armor platoons. One company team and one armor company were active units stationed at Fort Knox, KY. An ARNG armor platoon and an active company team were also observed during the refinement phase of the SIMUTA contract.

These nine units completed a total of 45 training tables with all units completing at least three training tables. Also, seven units began with a fundamental training table and two began with an offensive training table (see Table 1). Only two defensive training tables were executed by the units observed. Further information regarding the types of training tables completed by these units is presented in Appendix A.

Table 1

Number of Training Tables Executed by Observed Units

	Training Table						
Type of Training Table	1st	2nd	3rd	4th	5th	6th	7th
Fundamental	7[a]	7	4	1	0	0	0
Offensive	2	1	3	7	8	3	1
Defensive	0	1	1	0	0	0	0

[a] Number of units per training table.

Instrument. The RCVTP Training Observation Form was created to measure the different aspects of tactical skills expertise. This instrument, which can be found in Appendix B, contained the following sections:

1. Table preparation, which consisted of columns for recording the (a) time taken for preparation and (b) O/C coaching during preparation. (Reasons for sampling coaching are provided in the discussion of the performance measures.)

2. Events guide, which was created by the SIMUTA contractor team. This guide contained the prescribed sequence of O/Cs', exercise/controllers' (E/C) and units' actions for the training table. It also included a column for comments about the activity.

3. Addendum, which contained columns for noting additional unit activities and/or problems during a mission, such as radios not working or the unit's failure to send a report.

4. Post-exercise review of RCVTP training, which was comprised

of Likert-scale items regarding the following aspects of performance: (a) being disoriented, (b) following proper radio procedures, (c) keeping the O/C informed, (d) having problems with formations, and (d) losing control of the unit. The 1-5 scale used for these items--5 as always and 1 as never--was based on a scale developed by Kraemer and Wong (1992). These items were based on critical subtasks associated with the RCVTP.

5. AAR, which included columns for coded entries about the following aspects of the AAR: (a) agenda, (b) topics discussed, (c) communication processes (e.g., group discussion), and (d) time for each entry.

6. AAR summary items, which consisted of 22 Likert-scale items dealing with the O/C's and unit's activities during the AAR. Kraemer and Wong's (1992) scale was also used for these items.

The observational instrument also contained the following aids to help the observers: (a) instructions for completing the Table Observational Form and the AAR forms; (b) an annotated example of a completed Table Observational Form; and (c) a list of codes for AAR events. These aids are shown in Appendix C.

Training the observers. The observers'--two research psychologists and a graduate student intern with the Armored Forces Research Unit at Fort Knox, KY--training consisted of systematically going over a detailed set of instructions. They were also requested to view videotapes of an AAR conducted when the RCVTP training tables were being piloted. Also, two of the observers read the Army manual on platoon tactics (FM 17-15: U. S. Department of the Army, 1987). They discussed this manual with the Research and Development Coordinator for the Armored Forces Research Unit. Finally, problems or issues encountered during the first trial run were discussed by the observers. The third observer was quite knowledgeable with regards to platoon and company tactics.

Data collection procedure. Data were collected for a sampled unit by three observers. Because of constraints imposed by the training trial procedures, these observers were rarely able to record data for the same training tables. The sequence of training tables viewed by each observer varied from unit to unit. Observer A, for example, viewed the first three training tables for Unit 1 and the last three training tables for Unit 2. This variation helped to control for possible data biases due to systematic observer differences.

Each training table, as discussed, began with the preview. A typical preview took fifteen minutes. An additional amount of time, which varied with each training table and unit, was then spent at the simulators, preparing for the mission. The observers, correspondingly, recorded the information associated with the Table Preparation Form by following the unit activities at the O/C station and simulators. They also recorded any

downtime due to problems with the simulators (SIM).

The observers viewed the different units' execution of a training table from the stealth vehicle and the PVD at the appropriate O/C station. As indicated, they used the Table Observational Form to note the actions taken and problems encountered by the units, including any simulator downtime. Each occurrence of coaching provided by the O/C was also noted. The Post-exercise Review Form was completed immediately after the mission was over.

The platoon and company AARs began approximately fifteen and thirty minutes, respectively, after the training table was completed. The platoon AARs consisted of all platoon members, while the company AARs only included the company commander, his staff, platoon leaders, and tank commanders. The observers watched the appropriate AARs and recorded data from the back of the O/C's station. This viewing location made the discussion difficult to hear. Recording the topics being discussed was especially problematic.

A reliability check was conducted as observers were able to follow the same training tables for one unit. Few discrepancies were found among observers with regards to the performance data. Discrepancies were found, however, regarding the AAR data. (See Appendix D for an account of this reliability check.)

Criterion measures. As indicated, this assessment was comprised of performance and AAR data. The performance data consisted of:

1. Preparation time--minutes spent by units between the preview and Redcon 1 (ready to go) in preparing for the mission.

2. Exercise time--minutes elapsed from the O/C's first command to the last action for the table with SIM downtime excluded.

3. Errors--number of errors associated with conducting a formation/movement, navigating, and reporting to the O/C.

4. Coaching--number of coaching incidents recorded by the observers, including those during the preparation period. Some examples of coaching were prompting use of the radio and commanding units to report to the O/C. Skilled performance also involves the ability to attend to task cues without too much reliance on instructor prompting (Collins, Brown, & Newman, 1989; Patrick, 1992).

5. Table completion--determined by whether or not units performed all the critical subtasks associated with the training table.

6. Fratricides--determined by whether or not a friendly vehicle was killed during a training table. This measure was not

included with errors as it was not possible to record the number of fratricides per training table.

7. Post-exercise review--observer ratings on the post-exercise review items.

The AAR data consisted of:

1. Communication measures--separate proportions of the total AAR time found for O/C lecture, lecture by unit members, and group discussion.

2. Topic measures--separate proportions of AAR time spent talking about background issues, reasons for actions, and SIMNET issues.

3. AAR summary measures--observer ratings on the different AAR summary items and the total time for the AAR.

Scoring procedures. Two judges scored the different measures based on a predetermined scoring scheme. These judges--a research psychologist and graduate student intern at the Armored Forces Research Unit--also served as the observers.

Errors were scored based on guidelines for armor tactics as presented in FM 17-15. The AAR communication's data were scored with the stipulation that a "lecture" entry was only included with other like entries if it lasted a minute or more; otherwise, it was scored as discussion. All other measures were scored as indicated. The few discrepancies found in this scoring were resolved by a discussion between the judges.

Results and Discussion

Findings are discussed first for the performance measures and then for the AAR measures. See Appendix E[1] for a description of the statistical analyses associated with these measures.

Data for the performance measures. Trends in these units' exercise time, errors, and coaching scores were analyzed across successive RCVTP tables. Because of the limited sample size, these analyses also involved combining the data across platoons and companies and across active and ARNG units. Also, data for one training table were not recorded (see Table 2).

As shown in Table 2, units, typically, took less time, made fewer errors, and needed less coaching as their training progressed. Significant negative trends associated with these

---

[1]SPSS/PC+ procedures (Norusis, 1990) were used for the analyses done in this evaluation. The α-level for statistical tests of significance was set at a relatively liberal $p = .10$, in keeping with the exploratory nature of this evaluation.

measures were confirmed by statistical analyses (see Table E-1). These trends occurred even though the training tables were designed to become increasingly more difficult.

These trends were not a function of the units being less likely to finish their later training tables. As shown in Table 3, units were more likely to complete their fifth training table than their first training table. Also, the error data were not contaminated by the omission of the fratricide data as only one fratricide incident was reported for their last training table. The RCVTP thus seemingly helped these units to develop their collective tactical skills.

Questions exist, however, about the generalizability of the RCVTP's training effectiveness. These findings were based on a limited number of units and training tables. Only a few units, as stated, received any training regarding defensive exercises. Also, the raw data indicated that the units might have stopped improving after their third mission. If so, these units might then have only become proficient with regards to certain fundamental skills associated with tactics.

Trend analyses (See Table E-2) were then done for these units' exercise time, error rates and coaching scores for their offensive and defensive training tables. (These analyses were not done for the fundamental training tables because only two units completed three or more fundamental training tables.) As shown in Table 4, these units completed the later offensive and defensive tables significantly faster than they did the earlier ones. Units might then have become more adept during the course of their training to rapidly identify and destroy the enemy.

Table 2

Means and Standard Deviations of the Units' Time in Min, Error Rates, and Coaching Scores by Successive Training Tables

		Time in Min		Error Rate		Coaching Score	
Training Table	n	M	SD	M	SD	M	SD
First	9	85.22	30.40	12.89	4.81	8.44	5.27
Second	9	52.00	23.04	6.11	2.93	4.22	3.03
Third	8[a]	40.88	10.51	5.00	2.44	3.75	2.81
Fourth	8	41.00	12.59	6.38	4.43	5.25	3.99
Fifth	7	37.57	20.33	5.14	2.79	3.42	2.14
Sixth	3	32.00	10.44	1.67	.58	2.00	1.73

[a] Data missing for one unit.

Table 3

Exercise Completion Data by Successive Training Tables

	Training Table						
	1st	2nd	3rd	4th	5th	6th	7th
Number completed	4	6	7	7	7	2	0
Number not completed	5	3	1	1	1	1	1

The units' error rates and coaching scores also showed decreasing trends. However, these trends were not significant because some units experienced a reversal of the decreasing pattern for their fourth training table. This reversal might have been an artifact of the limited number of units sampled. Whatever the reason, the units did not get worse on the later training tables.

Table 4

Means and Standard Deviations of Units' Time in Min, Error Rates, and Coaching Scores by Successive Offensive/Defensive Training Tables

		Time in Min		Error Rates		Coaching Scores	
Training Table	n	M	SD	M	SD	M	SD
First	9	50.11	18.28	8.38	4.50	6.43	2.70
Second	9	31.55	8.16	4.67	3.10	2.80	2.93
Third	7	36.67	14.26	6.00	3.63	4.10	3.62
Fourth	4	28.50	5.32	4.00	4.39	3.50	1.92
Fifth	1	47.00	.00	4.00	.00	3.00	.00
Sixth	1	27.00	.00	2.00	.00	1.00	.00

Questions also remained as to whether the improvement trends cited previously were the results of RCVTP practice or other factors. The cited error and coaching trends might have been a result of these units' having more opportunities for making errors and being coached during their initial training tables. As shown, the sampled units took longer to complete their earlier training tables than their later ones. These completion times, however, might have been another indication of these units' having more problems with their earlier missions. Whatever the

case, further examination of the data indicated that the RCVTP was an effective training program, as the rates for errors and coaching by minutes were not higher for the later seemingly more difficult missions (see Appendix F).

The cited improvement trends could have also reflected units' becoming more adept at using the SIMNET system. These units did exhibit fewer problems with navigating in SIMNET on their second training table than on their first training table, with means of 4.00 and 1.56 for training tables 1 and 2, respectively. Also, the observers felt that units were more disoriented in their first training table than in the second training table, with reported means of 2.89 and 2.00 for Tables 1 and 2, respectively. This difference was statistically significant, $t(8) = 2.10$, $p < .10$. (See Appendixes G and H for these error data and post-exercises summary data.)

Finally, these units tended to spend less time in preparing for their later training tables than for their earlier ones (see Table 5). These data represented fewer sampling units than the previously cited data because the participants were not always given time to prepare for each mission and because preparing for a training table occasionally occurred when the observers were not able to observe, e.g., at night. Regardless, these data provided further evidence that these units were having less trouble preparing for a mission as their training progressed.

Data for the AAR measures. As indicated, these data were problematic. Poor reliability was found among observers for the communication process data. Also, observers had trouble following and recording the content of these discussions. Finally, fewer units were sampled for these data than for the performance data because AARs were not always given after each mission. The AAR data are summarized in Appendixes I-J.

Table 5

Means and Standard Deviations of Units' Preparation Time by Successive Training Tables

Training Table	n[a]	M	SD
First	9	29.00	11.29
Second	3	24.00	3.46
Third	7	24.85	14.62
Fourth	6	17.00	5.51
Fifth	4	18.75	.50
Sixth	3	9.67	8.50
Seventh	1	16.00	.00

[a] Numbers of units for this measure differ from other measures because of incomplete data.

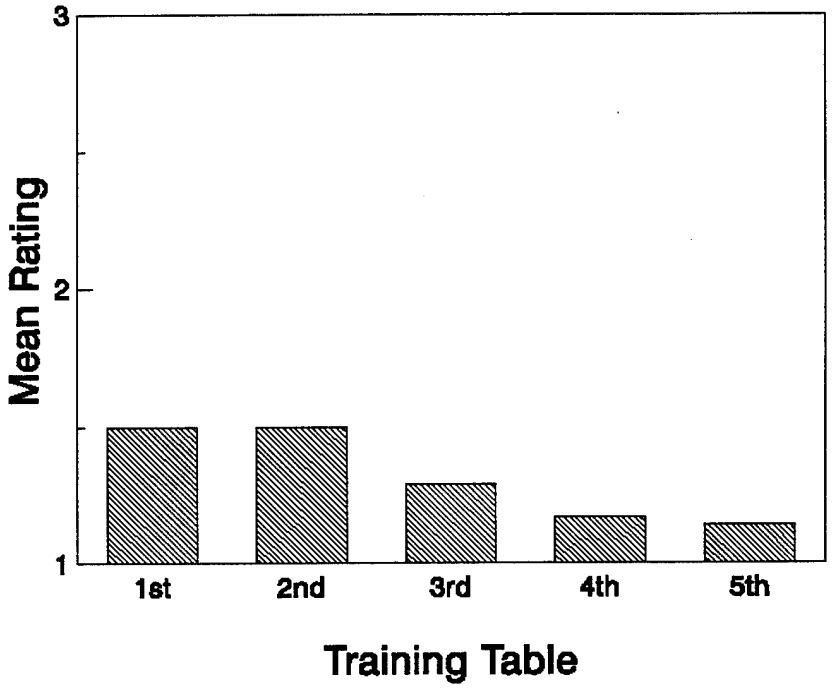


Figure 1. Observers' ratings of the proportion of participants' AAR comments dealing with SIMNET problems. Rating scale anchors include: 1 = NEVER (0%), 2 = SELDOM, 3 = APPROX. 50% OF THE TIME. Unit numbers at Orders 1-5 are 8, 8, 7, 6, 7, respectively.

The AAR data did indicate that SIMNET-related problems were not an issue for these units. Fewer than 1% of their reported comments in any given AAR dealt with SIMNET. Also, observers indicated that the participants rarely asked questions about or made comments about using SIMNET, with an overall mean rating of 1.30 for this AAR summary item (see Figure 1).

Also, these data showed that the typical AAR took nearly 30 minutes, which was the time suggested by the SIMUTA instructional design team for this activity. In addition, the units tended to need less time for an AAR as their training progressed (see Appendix H). Perhaps this finding is another indication that these units had fewer problems with their later training tables.

Summary of the observers' data. These data indicated that the RCVTP seemingly helped these units to develop their collective tactical skills expertise. That is, these units completed the later tables more quickly, with fewer errors, and with less coaching than the earlier ones.

However, additional evidence based on a larger sample is needed to confirm these findings. This evidence also needs to be based on performance judgments made by subject-matter experts. Such judgments are described in the next section, Assessment B.

Assessment B: The O/Cs' Judgments

This assessment was based on data found in the THPs regarding unit performance. Fourteen O/Cs provided these data. An O/C typically assessed the performance of four units with one being responsible for eight units. Occasionally, two or more O/Cs were identified as working together on an assessment.

Method

Participants. Data were collected on 38 armored force units. These units consisted of 17 armor platoons, 10 armor companies, 5 scout platoons, and 6 mechanized infantry platoons. Only 5 active units (4 armor companies and 1 scout platoon) were included in this sample.

These units executed 187 training tables, including 32 fundamental, 89 offensive, and 66 defensive training tables. Eighteen units began their training with a fundamental training table, 11 with an offensive table and 9 with a defensive table. Also, 89% of these units executed four or more training tables, 58% executed five or more training tables, and 45% executed six or more training tables. Table 6 shows the types of training tables completed by unit type.

Data collection procedure. For each training table completed by these units, the O/Cs indicated in the THP those subtasks which units needed either to "train to sustain" or "train to improve," representing satisfactory or unsatisfactory performance, respectively. These judgments were based on notes taken as units executed the missions.

Scoring procedures and measures. Two independent judges identified subtasks that these units performed at least twice. It was then possible to categorize the O/Cs' ratings into measures indicating changes in these units' subtasks proficiency as their training progressed. One set of measures dealt with subtask proficiency changes associated with units' initial and final performance of a subtask. These measures were:

1. Positive change, which consisted of the number of initial "train to improve" and final "train to sustain" ratings per subtask (i.e., improve/sustain).

2. No change, which consisted of the number of initial and final "train to sustain" ratings per subtask (i.e., sustain/sustain).

3. No change, which consisted of the number of initial and final "train to improve" ratings per subtask (i.e., improve/improve).

4. Negative change, which consisted of tabulating the number of initial "train to sustain" and final "train to improve" ratings per subtask (i.e., sustain/improve).

Also tabulated was the total number of "train to improve" and "train to sustain" ratings for these units' initial and final performance of the different subtasks.

Table 6

Number of Training Tables Executed by Unit Type

Type of Training Table	Training Table						
	1st	2nd	3rd	4th	5th	6th	7th-9th
Fundamental							
Armor Platoons	5	1	2	2	0	0	0
Mech/Scout Platoons	7	5	1	0	0	0	0
Armored Companies	6	2	1	0	0	0	0
Offensive							
Armor Platoons	5	9	8	6	2	3	0
Mech/Scout Platoons	3	4	4	4	3	2	2
Armored Companies	3	6	7	8	7	2	2
Defensive							
Armor Platoons	7	7	7	8	7	5	3
Mech/Scout Platoons	1	1	4	4	1	0	0
Armored Companies	1	2	2	2	2	5	0

Another set of measures involved examining these units' subtask proficiency across training tables. These measures consisted of counting, separately, the number of ratings for each training table which dealt with units': (a) first performance of a subtask (first subtasks) and (b) later performances of the same subtask (later subtasks). First and later subtasks were counted separately, because the former measure provided an indication of these units' subtask proficiency prior to the RCVTP. The first subtasks measure would also be indicative of these units' proficiency gained from other subtasks performed previously during this training. (See Appendix K for examples of the first and later subtasks which were found across training tables.)

All measures were aggregated by unit type. Percentages based on these aggregate counts were then computed. For the first and later subtask data, the percentage of tasks rated "train to sustain" was determined at successive training tables.

Results and Discussion

Findings are discussed first for the initial and final performance data. They are then discussed for the first and later subtasks data, which includes separate discussions of these

data across and by unit type. See Appendix L for a discussion of the statistical procedures used in this assessment.

Data for the initial and final performance measures. As shown in Table 7, a total of 359 subtasks had at least two ratings. Based on these frequencies, the percentage of subtasks with train to sustain ratings increased from 61.8% to 78.6 %. Furthermore, when the subtasks with ratings of train to improve were compared to the subtasks with ratings of train to sustain, a significant majority (74.6%) of them were train to sustain.

These units seemingly thus became more proficient on these subtasks as their training progressed. Data analyses confirmed this observation as significantly more subtasks were included in the improve/sustain category as compared to the subtasks included in the improve/improve category. Also, significantly more subtasks were included in the sustain/sustain category than in the sustain/improve category.

Similar comparisons were also done by unit type (see Table L-1). These comparisons revealed that the armor and mech/scout platoons were more likely to improve than were the armor companies. In subtasks rated train to sustain, for example, the armor companies' gain from first to last rating (60.5% to 71.1%) was about half that for the armor and mech/scout platoons (62.4% to 82.0%). However, this difference was not statistically confirmed.

Table 7

Initial and Final Subtask Rating Counts by Unit Type

		Positive Change	No Change		Negative Change
Type of unit	n	Improve-sustain	Sustain-sustain	Improve-improve	Sustain-improve
Armor Platoons	17	46	90	20	12
Mech/scout Platoons	11	20	45	6	6
Armored Companies	10	25	56	20	13
Total	38	91	191	46	31

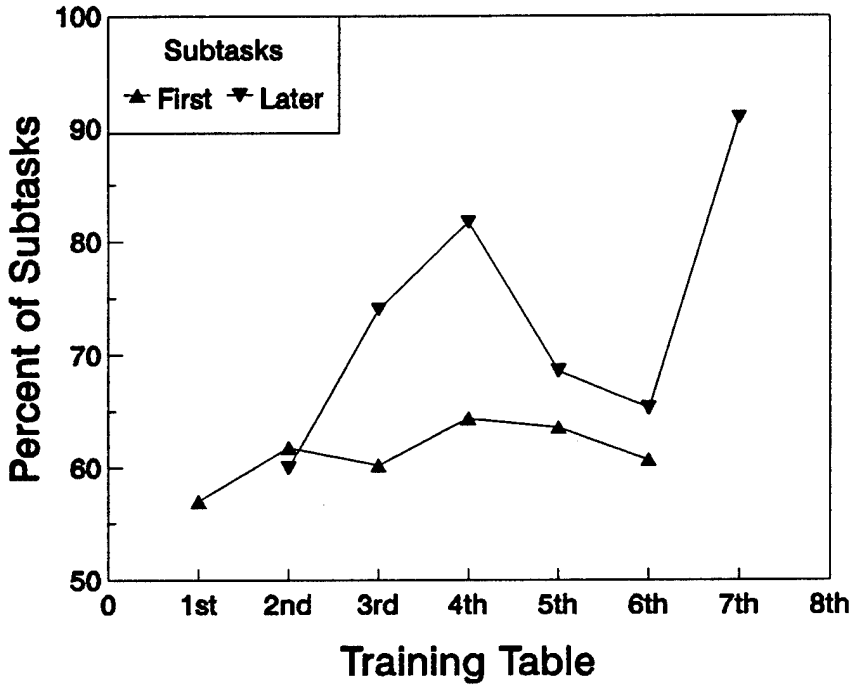


Figure 2. Percent of "first" and "later" subtasks with "train to sustain" ratings by successive training tables. (See Table L-2 for the number of subtasks and units at each point.)

Data for the first and later subtasks by unit type. As shown in Figure 2, the sustain percentage for these units first performance of a subtask varied around 59.9%, with only a modest increase across successive tables. This trend suggested that some small generalized transfer effects helped offset an expected decrease in performance when the more difficult subtasks were initially encountered in later training tables. As previously stated, these transfer effects were most likely associated with units' subtask proficiency prior to the RCVTP or gained from subtasks performed previously during this training.

Also, units' later performances on subtasks increased substantially for their third, fourth, and seventh training tables. Small increases were found for their fifth and sixth training tables. However, with only six data points, there was insufficient evidence to establish a systematic trend in units' improvement across training tables.

Further examination of these data did indicate that a significantly higher percentage of "train to sustain" ratings were recorded across units' second through sixth training tables for their first and later performances of the same subtasks. (See Table L-5). These units seemingly became more proficient as the result of practice afforded by the RCVTP.

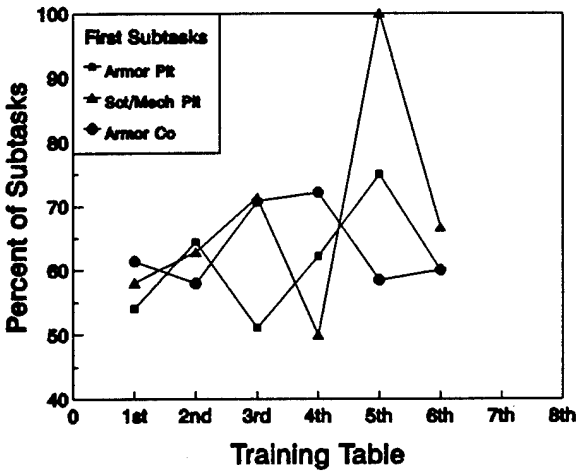


Figure 3. Percent of the "first" and "later" subtasks with "train to sustain" rating by successive training tables and unit type. (See Table L-2 for the number of subtasks and units at each point.)

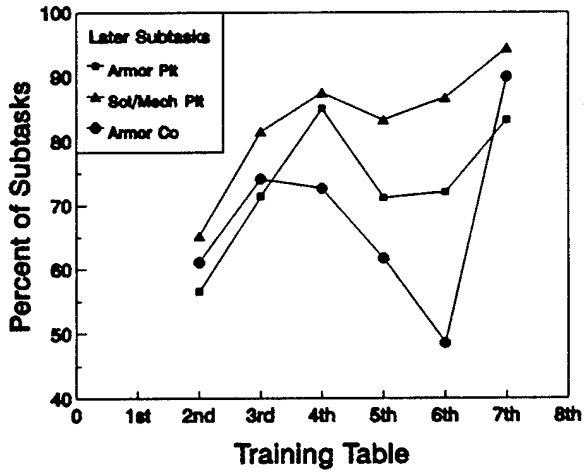


Figure 4. Percent of "later" subtasks with "train to sustain" ratings by successive training tables and unit type. (See Table L-2 for the number of subtasks and units at each point.)

Data for the first and later subtasks by unit type. Regardless of unit type, slight but inconsistent gains across training tables were found for these units' first performance of a subtask (see Figure 3). The pattern for the scout/mech platoons can be discounted, since they performed very few subtasks beyond the third training table.

Trends associated with the later performance data appeared to differ among units (see Figure 4). The decline in the fifth and sixth training tables was greater for the armor companies than for the platoons. This result provides further evidence showing that this training was more beneficial to the platoons than to the armor companies. This difference, however, was not verified by statistical analyses (see Table J-4).

Summary of O/Cs' Judgments. This assessment has provided additional evidence for the training effectiveness of the RCVTP. The O/Cs, who were subject-matter experts, indicated that this program did help a sizeable number of units to become more proficient in the critical subtasks associated with successful RCVTP performance. This improvement was not simply a function of learning to use SIMNET since final ratings for the different subtasks tended to be higher than those for the initial ratings. Also, platoons gained more from this program than did armor companies. Some answers have thus been provided to the questions regarding the observational data's validity. Questions, however, still remain about the participants' feelings toward this training program.

Assessment C: The Units' Questionnaire Responses

This section is based on the formative assessment of the RCVTP as conducted by the SIMUTA instructional design team.

Method

Participants. Questionnaire data were collected on 280 participants from the developmental trials. Two hundred thirty-nine of them were unit leaders, e.g., company commanders, platoon leaders, and tank commanders. The remaining participants were tank crewmen from armor companies. These participants came from: (a) 19 armor companies, (b) 12 armor platoons, (c) 3 scout platoons, and (d) 3 mechanized infantry platoons.

The participants per unit type were as follows: (a) 210 from armor companies; (b) 32 from armor platoons and (c) 38 from mechanized and scout platoons. Also, these participants included 206 ARNG and 74 active soldiers. The active soldiers were from armor companies.

Instrument and measures. The instrument used for this assessment was a questionnaire created by the SIMUTA instructional design team. The ARI evaluators also provided input regarding the questionnaire's design. Eight Likert-scaled items from this questionnaire were pertinent to the questions being examined in this assessment. These items dealt with:

1. Proficiency **before** the training--with 7 as extremely proficient, 1 as not at all proficient and four as a neutral point.

2. Proficiency **after** the training--which used the same scale as described for the first item.

3. Training benefits of the RCVTP compared to other SIMNET training experience--with 7 as extremely high, 1 as extremely low, and 4 as the neutral point.

4. Improvement due to simulator time--with 7 as very much and 1 as not at all.

5. Improvement due to AARs--which used the same scale as cited for the fourth item.

6. Training tables went from less to more difficult--with 7 as strongly agreeing with that statement, 1 as strongly disagreeing, and 4 as a neutral point. (This scale was used for the remaining items.)

7-8. The participants' comments during the AARs helped the unit to improve (an item for the company and platoon AARs).

The first two items listed were designed to provide insights

into participants' perceptions regarding changes in their unit's proficiency during the course of this training. Item 4 dealt with their perceptions about the relative training benefits of this program. Data from the remaining items were expected to provide insights into the participants' feelings about reasons for these benefits and different aspects of the RCVTP.

Data collection procedure. All ethical guidelines prescribed by ARI and the American Psychological Association were followed when the questionnaire was administered at the end of the participants' training. Participants were told that the responses were to be used for research purposes only and that their names were not to be included on the questionnaire.

Results and Discussion

Two caveats must be made before presenting and discussing these results. One, the n fluctuated among the selected items because: (a) the SIMUTA team gave 88 participants an abbreviated form of the questionnaire to complete, (b) some participants failed to answer questions, and (c) some items (e.g., those on company AARs) were not germane to all participants. Two, the data for items dealing with the relative training benefits of the RCVTP and reasons for any training benefits were collapsed across the sample. This was done because these evaluators were only looking for general trends in these measures.

The number of respondents, mean and standard deviation for each item are presented in Appendix M. The statistical analyses associated with this assessment are discussed in Appendix N.

Data regarding changes in perceived levels of proficiency. Significant differences for all three types of units were found regarding estimates of pre- and post-RCVTP training proficiency levels. The participants, regardless of unit type, claimed to be more proficient after training than they were before training (see Figure 5). Also, comparisons between unit types on a difference score (before training estimates minus post-training estimates) revealed a "significant interaction" between unit types and proficiency-levels. As shown in Figure 5, the leaders from the armor platoons and mech/scout platoons indicated more improvement than did their counterparts in the armor companies. This interaction was primarily a function of the platoons claiming to be at a lower level of initial proficiency than the participants from the armor companies.

As shown in Figure 6, leaders from the reserve company units indicated that they improved more than did their active counterparts. This interaction seemingly occurred because the ARNG unit leaders claimed to be at a lower level of initial proficiency than did their active counterparts. This training program thus raised the confidence levels of the ARNG armor company leaders to the claimed pre-training levels of the active company leaders.

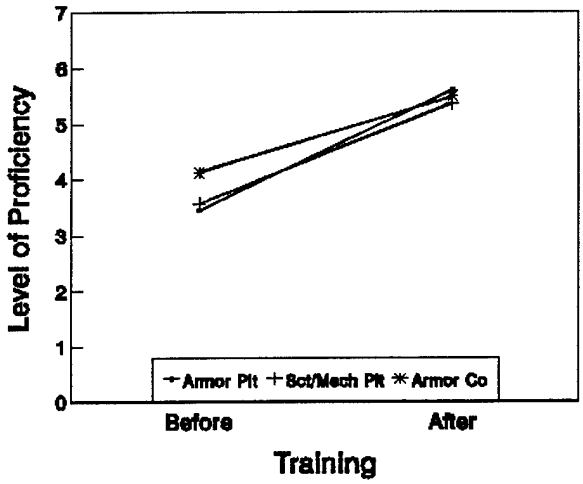


Figure 5. Means of proficiency estimates by unit leaders from armor platoons, mech/scout platoons, and armor companies.

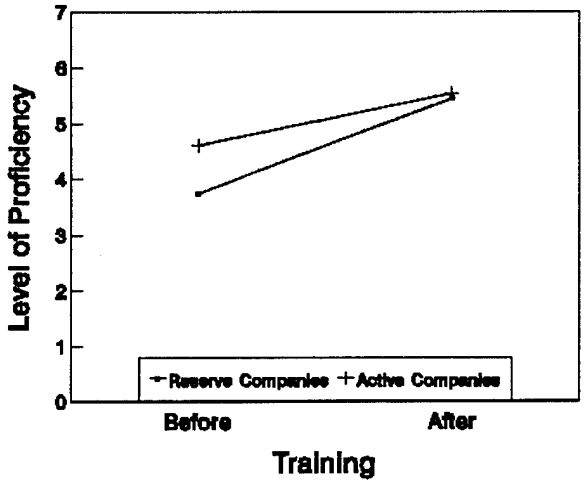


Figure 6. Means of proficiency estimates by unit leaders from reserve component and active armor companies.

Data regarding the participants' perceptions of the RCVTP. The participants, regardless of their designation, believed that improvement in their unit's performance was a function of the RCVTP. Means of 5.43 and 5.54 were found for the items dealing, respectively, with improvement as a function of the time in the simulators and the AARs. They, furthermore, indicated that they became more proficient after this training than after their other SIMNET training experiences with a mean of 5.66 for this item.

The questionnaire data also provided some insights into the participants' feelings about components of the RCVTP's instructional design. One, they felt that discovery learning did take place. This point was manifested by their scores for the items dealing with their AAR comments with means of M=5.75 for platoon training tables and 5.70 for company training tables. Two, these training tables were viewed as becoming more difficult as the training progressed. A mean of 5.55 was found for the item dealing with this issue.

Summary of units' questionnaire responses. Questions have then been answered about the participants' feelings about this program. These participants, especially those from the platoons and the ARNG armor companies, indicated that the RCVTP helped them and correspondingly their units to become more tactically proficient. Evidence has also been obtained indicating this program helped ARNG participants feel that they were as tactically proficient as their counterparts from the active armor companies. This program's effectiveness has thus been established from the perspective of ARNG users. ARNG units seemingly then would like to utilize this program for their future collective tactical training.

Summary and Conclusions

Data from the different methods and sources indicate that the units developed their collective tactical skills across the training period. However, as noted, the amount of practice associated with the RCVTP was apparently adequate for platoon-level training, while armor companies might have needed some additional practice on some tables. This evaluation has thus demonstrated the RCVTP's instructional value for helping tactical units--especially platoons--to become more proficient. Correspondingly then, this instructional program can apparently provide ARNG soldiers with some of the training necessary to handle their new roles in the post cold-war military.

This evaluation has also indicated that some aspects of the RCVTP were utilized as designed while others were not. As discussed, the tables were perceived by the sampled units to become progressively more difficult. However, the units were observed to spend more time in preparing for missions but less time in executing them than intended.

This evaluation has finally demonstrated the value of employing a multimethod-multisource evaluation strategy for conducting naturalistic evaluations of high technology based training programs. As stated, each method might have provided problematic data. The observational data, for example, were limited by their small sample size and the exclusion of defensive training tables. Areas of agreement among assessments thus provided more valid conclusions than any single assessment method would have provided.

Also, each assessment yielded insights into this training situation from complementary perspectives. As indicated, the observational data reflected the perspective of evaluators who were independent of the instructional design and training processes; however, they were not subject-matter experts. The instructors who were subject-matter experts were also part of the instructional process. The questionnaires tapped the users' perspectives. Taken from these different methods and sources, the evidence for RCVTP's effectiveness becomes more compelling.

However, limitations with this evaluation have precluded the possibility of making any definitive conclusions about the RCVTP's effectiveness for training ARNG units. These evaluators have not been able to collect data on a control group of ARNG units who went through SIMNET training but not the RCVTP. Perhaps, significant differences would not have been found in such a comparison. If so, the RCVTP then might have limited training benefits for ARNG units as such units would benefit just as much from "non-structured" SIMNET-based practice. Future evaluations of this system must involve, if possible, comparisons between ARNG units receiving RCVTP and an equivalent control group.

In conclusion, the following recommendations can be drawn from this evaluation:

The RCVTP should continue to be used to train ARNG units.

A "multimethod-multisource" approach should be used, when possible, to conduct naturalistic evaluations of high technology training systems.

Further research is needed to provide more definitive conclusions regarding the RCVTP's effectiveness for training ARNG and other units.

References

Bessemer, D. W. (1991). Transfer of SIMNET training in the Armor Officer Basic Course (ARI Technical Report 920). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A233 198)

Brown, R. E., Pishel, R. E., & Southard, L. D. (1988, April). Simulation networking (SIMNET) preliminary training developments study (PTDS) (TRAC-WSMR-TEA-8-88). (Available from the U.S. Army Armor School, ATTN: ATSB-DOTD-ORA, Fort Knox, Ky 41021-5200.)

Burnside, B. L. (1982). Subjective appraisal as a feedback tool (ARI Technical Report 604). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A138 873)

Campbell, C. H., Campbell, R. C., Sanders, J. J., & Flynn, M. C. (In preparation). Methodology for development of structured simulation-based training (ARI Research Product). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Collins, A., Brown, J. S., & Newman, S. E. (1989). Cognitive apprenticeship: Teaching the crafts of reading, writing, and mathematics. In L. S. Resnick (Ed.) Knowing, learning and instruction: Essay in honor of Robert Glaser (pp. 453-459). Hillsdale, NJ: Erlbaum.

Cook, T. D., & Campbell, D. T. (1979). Quasi-experimentation: Design and analyses issues for field settings. Chicago: Rand McNally.

Denizen, N. K. (1978). The logic of naturalistic inquiry. In N. K. Denizen (Ed.) Sociological methods: A sourcebook. New York: McGraw-Hill.

Garvey, E. E., & Radgowski, T. (1988). SIMNET-D standing operating procedure (Report No. 6929). Cambridge, MA: BBN Systems and Technologies Corporation.

Gound, D., & Schwab, J. (1988, March). Concept evaluation program of simulation networking (SIMNET) (Final Report, TRADOC TRMS No. 86-CEP-0345, ACN 86299). Fort Knox, KY: U.S. Army and Engineer Board. (Available from Commander, U.S. Army Armor Center and Fort Knox, ATTN: ATZK-DS, Fort Knox, KY 40121-5180).

Herrmann, D. J. (1982). Know thy memory: The use of questionnaires to assess and study memory. Psychological Bulletin, 92, 343-352.

Hoffman, R. G., Graves, C. R., & Koger, M. E. (In preparation). Developing the RCVTP: History and lessons learned (ARI Research Product). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Kendall, M. K. (1975). Rank correlation methods (4th Ed.). London: Charles Griffin & Company, Ltd.

Kraemer, R. E., & Bessemer, D. W. (1987). U.S. tank platoon training for the 1987 Canadian Army Trophy (CAT) competition using a simulation networking (SIMNET) system (ARI Research Report 1457). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A191 076)

Kraemer, R. E., & Wong, D. T. (1992). Evaluation of a prototype platoon gunnery trainer for Armor Officer Basic Course training (ARI Research Report 1620). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A254 289)

Meliza, L. L., Bessemer, D. W., & Tan, S. C. (1994). Unit Performance Assessment System Development (ARI Technical Report 1008). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD B191 576)

Morrison, J. E., & Holding, D. H. (1990). Designing a gunnery training strategy (ARI Technical Report 899). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A226 129)

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than one can know. Verbal reports on mental processes. Psychological Review, 84. 231-259.

Norusis, M. J. (1990). SPSS/PC + 4.0 base manual for the IBM PC/XT/AT and PS/2. Chicago: SPSS Inc.

Patrick, J. (1992). Training: Research and practice. London: Harcourt Brace Jovanovich, Publishers.

Patton, M. Q. (1987). How to use qualitative methods in evaluation. Newbury, CA: Sage Publications, Inc.

Shlechter, T. M., Bessemer, D. W., & Kolosh, K. P. (1991). Role-playing and the effectiveness of a high technology training simulator (ARI Technical Report 938). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A224 913)

Shlechter, T. M., Bessemer, D. W., Rowatt, W. C., & Nesselroade, K. P., Jr. (1994). Evaluating the Unit Performance Assessment System's after action review displays (ARI Technical Report 997). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A281 712)

Shute, V. J., & Regian, J. W. (1993). Principles for evaluating intelligent tutoring systems. Journal of Artificial Intelligence and Education, 4, 245-271.

Turcek, J. L., Campbell, C. H., Myers, W. E., & Garth, T. H. (In preparation). Reserve Component Virtual Training Program (RCVTP) orientation guide (ARI Research Product). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

U.S. Department of the Army. (1987, October). Tank platoon (FM 17-15). Washington, DC. (Available from the Commander, U.S. Army Armor Center and Fort Knox, ATTN: ATZK-DS, Fort Knox, KY 40121-5200)

Appendix A

List of the RCVTP Tables Sampled in Assessment A

Types of Tables Executed By the Units

Table	Table Description	Number of Units per Table
Armor Platoons		
Fundamental Training Tables		
PAA1	Basic Movement Skills	2
PAA2	Tactical Movement; Actions on Contact	2
PAA3	Basic Defensive Techniques	1
Offensive Training Tables		
PAB1	Tactical Road March	2
PAB2	Movement into Battle; First Contact	1
PAB3	Continued Movement; Platoon Reacts to Contact	1
Mechanized Platoons		
Fundamental Training Tables		
PMA1	Basic Movement Skills: Command & Control	1
PMA2	Tactical Movements; Actions on Contact	2
PMA3	Basic Defensive Techniques	1
Offensive Training Tables		
PMB1	Tactical Road March	1
PMB2	Tactical Movement; Initial Contact	2
PMB3	Continued Movement; Platoon Reacts to Contact	2
Armor Company/Team		
Fundamental Training Tables		
CAA1/CTA1	Tactical Road March; Command and Control	2
CAA2/CTA2	Basic Tactical Movement Skills; Actions on Contact	3
Offensive Training Tables		
CAB1/CTB1	Tactical Road March	2
CAB2/CTB2	Tactical Movement; Initial Contact	4
CAB3/CTB3	Continued Movement; Company Reacts to Contact	3
CAC1/CTC1 CAC2	Mission Changed; Increased Contact	2
Defensive Training Tables		
CAF1/CTF1	Defense	1
CAF2/CTF2	High Risk Defense	1

Appendix B

A Prototype Observational Form

SIMNET TRAINING OBSERVATION FORM:

TABLE PREPARATION

Training Observer: _____ Date: _____

Type of Unit: _____ Time: _____

TIME	COACHING	UNIT ACTIVITIES	NOTES/COMMENTS

COACHING:

COM = COMMANDING

PROP = PROMPTING

OTHER--SPECIFY

SIMNET TRAINING OBSERVATION FORM

ADDENDUM

Training Observer: _____

Date: _____

Time: _____

Training Instructor:_____

Organization: _____

Time scheduled to reach LD/SP: _____

INSTRUCTIONS: When observing an exercise several events may occur. As the observer you should note the time, activity, and persons participating in major events. For this form, note any problems or activities encountered which could not have been noted on the O/C list of activities.

COMMENTS:_____

_____

_____

_____

_____

TIME	ACTIVITY	COMMENTS

INSTRUCTIONS FOR OBSERVING SIMUTA TRAINING

Training Observer: _____ Date: _____

Type of Unit: _____ Time: _____

INSTRUCTIONS: **AS SHOWN IN THE ANNOTATIVE EXAMPLE, OBSERVATIONS ARE TO BE RECORDED ON THE TABLE EVENT LIST. OBSERVERS WILL THUS NEED A COPY OF THE SPECIFIC TABLE EVENT LIST FOR THE EXERCISE BEING OBSERVED.**

When observing the exercise, indicate the **TIME** when the listed action took place. Time should be indicated to the **right of the action.** Put a dash by any platoon action which occurs within one minute of the OCIC action; otherwise, indicate the time for the platoon action.

Also note the type of coaching--**COMMANDING or PROMPTING**--that the unit received from either the O/C or E/C. **COMMANDING** refers to direct coaching by the O/C. It is illustrated when the O/C tells the unit to move out. **PROMPTING** refers to indirect coaching by the O/C, e.g., providing the unit with information that would make them move out. It also refers to actions (e.g. indirect fire) initiated by the E/C. **Prompting** differs from cuing in that the latter is not part of the scenario's script. **YOUR NOTATION REGARDING COACHING BELONGS IN THE COLUMN DESCRIBING THE OCIC's or E/C's ACTION.** Additional comments on actions taken by the OCIC, E/C or the corresponding PLT actions belong in the comment's column.

Finally, note the reasons for ending the exercise (e.g., mission accomplished) and any problems that the unit or O/C had with using the SIMNET system (e.g., a simulator crashed).

REASONS FOR ENDING THE EXERCISE _____

_____

_____

_____

ATTENTION! DON'T FORGET TO COMPLETE THE POST-EXERCISE REVIEW FORM

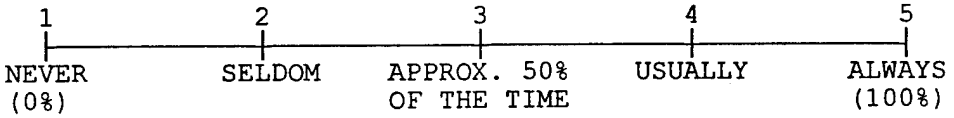
POST-EXERCISE REVIEW OF

SIMUTA TRAINING

Training Observer: _____ Date: _____

Type of Unit: _____ Time: _____

For each statement mark the line in the location which you feel best represents the units performance for the exercise. Use the following scale:



1. The unit was disoriented during the exercise. 1 2 3 4 5

2. The unit followed proper RTO procedures. 1 2 3 4 5

3. The unit kept the O/C informed. 1 2 3 4 5

4. The unit had problems with their formation. 1 2 3 4 5

5. The unit leader lost control of the unit. 1 2 3 4 5

OTHER OBSERVATIONS ABOUT THE UNIT'S PERFORMANCE:

_____

_____

_____

_____

SIMNET TRAINING OBSERVATION FORM:

AFTER ACTION REVIEW

Training Observer: _____ Date: _____

Type of Unit: _____ Time: _____

Location of AAR: _____

INSTRUCTIONS: To properly code the AAR, indicate the time and reference the sub-item each time the TOPIC, or communication PROCESS changes and also note the time when visual AIDS are employed. The contents of the topics discussed can be noted in the NOTES/COMMENTS column. The AGENDA ITEM, TOPIC and PROCESS need only be noted when they change categories.

Use the following scheme to code the AGENDA ITEM entries:

1 = REVISIT COMBAT FUNC'S/KEY TP'S
2A = LEADERS' PLAN
2B = ENEMY PERSPECTIVE
2C = BATTLEFIELD EXECUTION SUMMARY
3 = DISCUSSION
4 = REVIEW OF TABLE
5 = PREVIEW

PLEASE NOTE THAT PLATOON AARs DO NOT FOLLOW THIS AGENDA. FOR THOSE AARs, THE AGENDA WILL BE LISTED ON THE POSTER BOARD.

Use the following scheme to code the TOPIC entries:

BKG = BKG OF MISSION SIT
EVT = WHAT HAPPENED
ANL = WHY HAPPENED
IMP = HOW TO IMPROVE
SIM = SIMNET FACTORS
OTH = OTHER

Use the following scheme to code the communication PROCESSES entries (The PA entry is a sub-entry within the discussion entry. The number of PA's should be tallied in the space following PA:

DIS = DISCUSSION (all interaction e.g. questions/answers)
OLT = O/C LECTURE
ASU = AIDS SET-UP
SLT = STUDENT LECTURE
PA = PROMPTING ANSWERS
CL = CLARIFYING ANSWERS

PA occurs when the O/C encourages a reluctant student to respond. CL occurs when the O/C clarifies a student's comment/answer. Or when the O/C clarifies a question.

Use the following scheme to code the information regarding the O/C's use of AIDS:

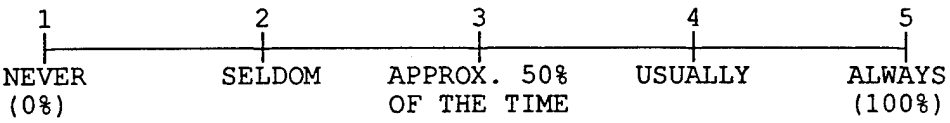
MAP = TERRAIN MAP
POS = POSTER
STL = STEALTH
UPAS= UPAS
WB = WRITING BOARD
OTH = OTHER

TIME	TOPICS	PROCESSES	AIDS	NOTES/COMMENTS
ADG___	BKG IMP EVT SIM ANL OTH	DIS OLT SLT ASU PA: CL:	MAP POS STL UPAS WB OTH	
ADG___	BKG IMP EVT SIM ANL OTH	DIS OLT SLT ASU PA: CL:	MAP POS STL UPAS WB OTH	
ADG___	BKG IMP EVT SIM ANL OTH	DIS OLT SLT ASU PA: CL:	MAP POS STL UPAS WB OTH	
ADG___	BKG IMP EVT SIM ANL OTH	DIS OLT SLT ASU PA: CL:	MAP POS STL UPAS WB OTH	
ADG___	BKG IMP EVT SIM ANL OTH	DIS OLT SLT ASU PA: CL:	MAP POS STL UPAS WB OTH	
ADG___	BKG IMP EVT SIM ANL OTH	DIS OLT SLT ASU PA: CL:	MAP POS STL UPAS WB OTH	
ADG___	BKG IMP EVT SIM ANL OTH	DIS OLT SLT ASU PA: CL:	MAP POS STL UPAS WB OTH	
ADG___	BKG IMP EVT SIM ANL OTH	DIS OLT SLT ASU PA: CL:	MAP POS STL UPAS WB OTH	
ADG___	BKG IMP EVT SIM ANL OTH	DIS OLT SLT ASU PA: CL:	MAP POS STL UPAS WB OTH	
ADG___	BKG IMP EVT SIM ANL OTH	DIS OLT SLT ASU PA: CL:	MAP POS STL UPAS WB OTH	
ADG___	BKG IMP EVT SIM ANL OTH	DIS OLT SLT ASU PA: CL:	MAP POS STL UPAS WB OTH	

General Comments:____

____

SUMMARY OF THE AFTER ACTION REVIEW: Mark the line in the appropriate location for each statement using the following rating scale:



QUESTIONS ABOUT THE AAR LEADER'S COMMENTS. For the following questions, mark the line in the appropriate location for each statement **in relationship to the number of comments made by the AAR leader.**

1. He acted as a facilitator of the group's discussion. 1 2 3 4 5
2. He stated the reasons for the group's actions. 1 2 3 4 5
3. He stated possible alternative actions. 1 2 3 4 5
4. He stated the lessons learned. 1 2 3 4 5
5. He discussed the skills to be sustained. 1 2 3 4 5
6. He matched tactical events to key teaching points. 1 2 3 4 5
7. He made critical comments of the groups's performance. 1 2 3 4 5
8. He interrupted the students' responses. 1 2 3 4 5
9. He posed questions just to the student leader(s). 1 2 3 4 5
10. He lost control of the group. 1 2 3 4 5
11. He lost time using visual aids. 1 2 3 4 5

List the types of visual aids which cost time: _____

QUESTIONS ABOUT STUDENT COMMENTS. For the following questions, mark the line in the appropriate location for each statement **in relationship to the number of comments made by the students.**

1. They stated the reasons for their actions. 1 2 3 4 5
2. They discussed possible alternative actions. 1 2 3 4 5
3. They stated the lessons learned. 1 2 3 4 5
4. They stated the skills to be sustained. 1 2 3 4 5
5. They made critical comments about the group's performance. 1 2 3 4 5
6. Their comments were prompted by the AAR leader. 1 2 3 4 5
7. Their comments were clarified by the AAR leader. 1 2 3 4 5
8. They asked questions about or made comments about using SIMNET. 1 2 3 4 5
9. Students other than the unit leader(s) participated in the AAR. 1 2 3 4 5
10. They used visual aids. 1 2 3 4 5

List the types of visual aids used: _____

FINAL OBSERVATIONS ABOUT THE AAR:

_____

_____

_____

Appendix C

Observational Aids

CHEAT SHEET FOR OBSERVERS

AGENDA ITEMS:

1 = REVISIT COM FUNC/KEY TP'S	3 = DISCUSSION
2A = LEADERS' PLAN	4 = REVIEW OF TABLE
2B = ENEMY PERSPECTIVE	5 = PREVIEW
2C = BATTLEFIELD EXEC SUMMARY	

EVENTS:

INT = INTRODUCTORY MATERIALS	ACT = ACTIONS ON CONTACT
MIS = MISSION/INTENT/SIT/TLP	SUM = SUMMARY OF THE BATTLE
MF = MOVEMENTS AND FORMATIONS	LL = LESSONS LEARNED
REP = REPORTS	SS = SKILLS TO BE SUSTAINED
RTO = RADIO TRANS OPERATIONS	SIM = SIMNET PROBLEMS/QUESTIONS
CUE = CUES/KEY EVENTS	OTH = OTHER

(O/C) PROCESSES:

AQ = ASKING QUESTIONS	PA = PROMPTING ANSWERS
LSL = LISTENING TO STUDENT LEADERS	RQ = RESPONDING TO QUESTIONS
LOS = LISTENING TO OTHER STUDENTS	DS = DISPLAY SET-UP
LT = LECTURING	OTH = OTHER

AAR AIDS:

MAP = TERRAIN MAP	UPAS = UPAS
PVD = PLAN VIEW DISPLAY	WB = WHITE BOARD/BUTCHER BOARD
STL = STEALTH	OTH = OTHER

Appendix D

Results of the Reliability Check

Inter-Observer Reliability (exercise CAC3 @14:30 3/11/94)

Entries	TS (observer 1)	DB (observer 2)
OCIC 1	14:45	14:45
EC 1	14:48	14:48
CO 1	14:58	14:53
CO 2	14:57	14:55
OCIC 2	15:04	15:02
EC 2	--	--
CO 3	--	--
OCIC 3	15:07	15:06
EC 3	--	--
CO 4	15:10	15:07
EC 4	--	--
CO 5	15:21	--
OCIC 4	15:19	15:18
CO 6	--	--

Post Exercise

1	2.75	3.00
2	4.75	4.00
3	3.50	3.00
4	2.75	3.00
5	2.25	4.00

Coaching (total)	2	1

AAR

Total Time	33	36
Lecture Time	14	13
Discuss Time	12/18*	19

*depending on coding rule-of-thumb

Inter-Observer Reliability (exercise CAC2 @12:30 3/11/94)

Entries	TS (observer 1)	PN (observer 2)
OCIC 1	12:48	12:48
EC 1	12:48	12:48
CO 1	12:49	12:48
EC 2	12:52	12:52
CO 2	12:53	12:52/12:53
OCIC 2	12:55	12:54
EC 3	12:54	12:53
CO 3	12:55	12:54
OCIC 3	13:00	13:00
OCIC 4	13:08	--
CO 4	13:07/13:10	13:06/13:10
OCIC 5	13:20	13:20
Post Exercise		
1	2.00	2.00
2	4.50	4.00
3	4.50	5.00
4	2.25	3.00
5	2.00	1.50
Coaching (total)	1	2
AAR		
Total time	35	35
Lecture Time	8	4
Discuss Time	22	27

Appendix E

Description of the Trend Analyses For Assessment A

Kendall's (1975) τ rank-order correlations were computed to determine the existence of any significant trends in the units' exercise time, errors, and coaching scores across successive exercises in their RCVTP training. These analyses consisted of ranking the measured values in relationship order in which each unit completed their training tables. If, for example, a unit completed exercises in 35 min, 40 min, and 45 min for their three training tables, respectively, then the third table received a ranking of 1 for having the largest time value. Based on these individual unit rankings, τ correlations were computed for the different measures. When a unit had equal values (and thus tied ranks) for two or more tables, then the correction for ties was used (Kendall, 1975).

Mean τ values were computed weighing the individual unit τs inversely by their estimated standard errors. Weighted two-tailed *t*-tests were then computed to determine if the mean τs differed significantly from zero. The α-level for these tests was set at $p = .10$. The results of the analyses are shown in Table E-1. Significant negative trends are found for all three variables.

Table E-1

Weighted Mean τ-Values and Tests of Significance Including All RCVTP Tables

Variables	M [a]	SD	df	t
Times	-.574	.274	8	6.31**
Errors	-.340	.174	8	5.86**
Coaching	-.206	.186	8	3.32**

[a] negative number indicates a decreasing trend.
** $P < .05$

Kendall's (1975) τs also were calculated for the units' exercise time, error rates and coaching scores just for their offensive and defensive training tables, disregarding the fundamental training tables. (Taus were not calculated for the fundamental training tables because only four units completed as many as three of these tables.) Two units that completed fewer than three offensive and defensive training tables could not be included in the analyses. The results of the analyses are shown in Table E-2. In this case the trend is significant only for the exercise time variable. Inspection of the unit τs for the error and coaching variables revealed varied trends. Larger values at the third table for some units disturbed the general pattern of

decreasing values.

Table E-2

Weighted Mean τ-Values and Tests of Significance for the RCVTP Offensive/Defensive Tables, Omitting the Fundamental Tables

Variables	M [a]	SD	df	t
Times	-.359	.280	6	3.39**
Errors	-.010	.470	6	.06
Coaching	.014	.237	6	.16

[a] negative number indicates a decreasing trend.
** P < .05

Appendix F

Results of Error Rates and Coaching Occurrences Per Unit Time

Additional data analyses of Assessment A's performance data have been conducted with errors and coaching occurrences divided by time in mins. These analyses have provided indexes of errors per unit time and coaching per unit time, which may also be considered as indexes of the relationship between occurrences of and opportunities for making errors and receiving coaching. As shown in Figures F-1 and F-2, decreasing trends have not been found in these data. However, as discussed in the text, the meaning of these results are not clear.

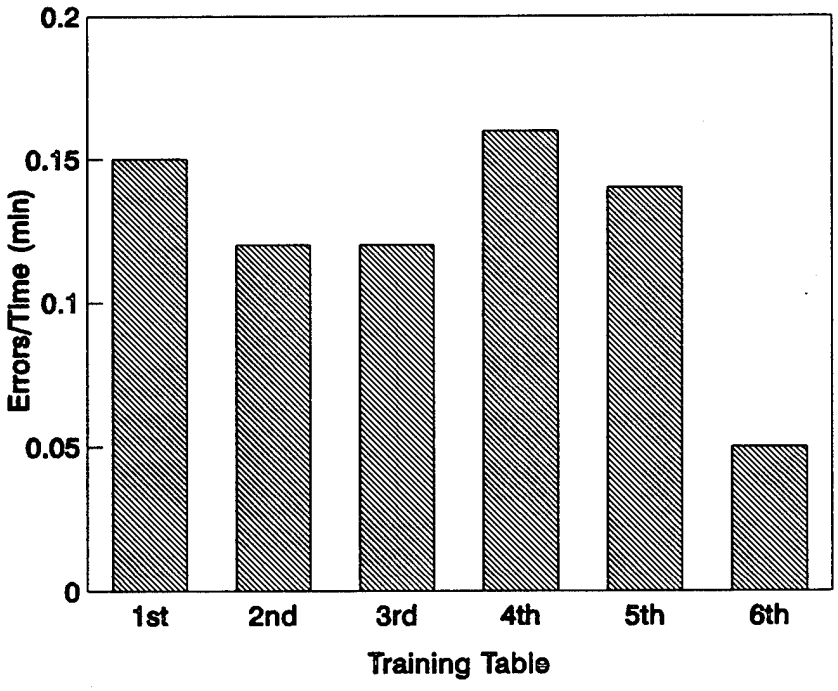


Figure F-1. Error rates (frequency of errors per unit time in min) for successive training tables.

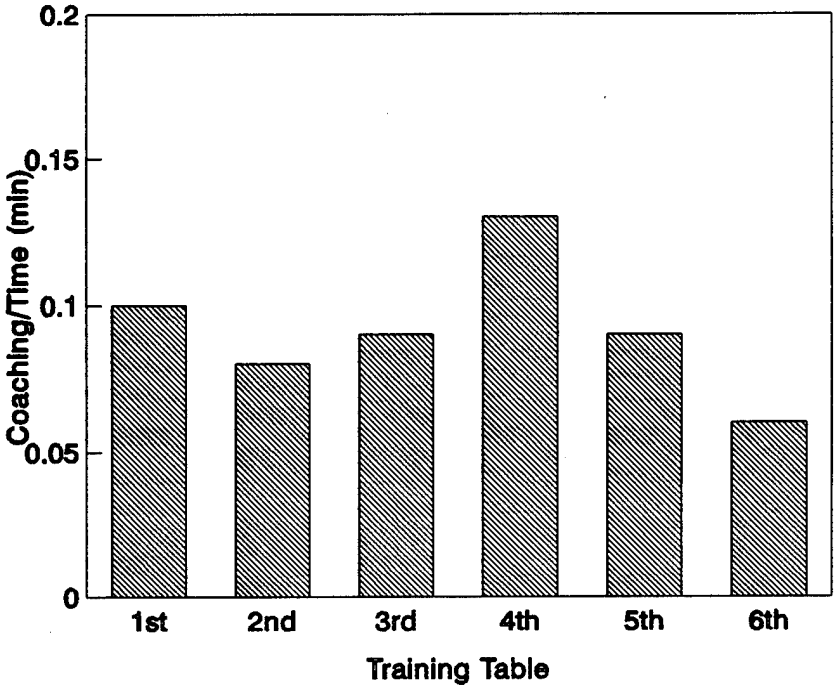


Figure F-2. Coaching rates (frequency of coaching per unit time in min) for successive training tables.

Appendix G

Results of the Error Data Broken Down by Component Elements

Table G-1

Means and Standard Deviations of the Units' Navigational, Formation, and Reporting Errors for Successive Training Tables

	Navigational Errors			Formation Errors		Reporting Errors		Other Errors	
Training Table	n	M	SD	M	SD	M	SD	M	SD
First	9	4.00	3.08	5.00	2.12	2.78	2.17	1.11	1.29
Second	9	1.56	1.42	1.89	1.53	2.11	1.17	.56	.53
Third	8	1.50	1.07	1.25	.88	1.38	1.30	.88	.88
Fourth	7	.63	.74	1.25	1.39	3.25	2.96	1.25	1.25
Fifth	8	1.57	2.14	1.14	1.35	2.00	.82	.53	.43
Sixth	3	.33	.58	.33	.58	.33	.58	.58	.67
Seventh	1	3.00	.00	1.00	.00	.00	.00	.00	.00

Appendix H

Results of the Post-Exercise Data

Table H-1

Means and Standard Deviations of the Post-Exercise Data for Successive Training Tables

		UNIT DISORIENTED		USED PROPER RADIO PROCEDURES		KEPT O/C INFORMED		PROBLEMS WITH FORMATION		LEADER LOST CONTROL	
Training Table	N[a]	M	SD	M	SD	M	SD	M	SD	M	SD
First	9	2.89	.93	3.56	1.13	3.22	1.20	3.00	1.22	2.44	1.01
Second	9	2.00	1.00	4.11	.78	3.44	.73	2.89	.93	1.89	.60
Third	8	2.63	1.06	3.25	1.55	3.25	1.28	2.75	.71	2.00	.93
Fourth	8	2.25	1.03	3.63	1.19	3.38	1.06	3.38	.74	2.25	1.04
Fifth	7	1.85	1.33	4.00	.82	3.43	.78	1.72	.95	1.85	1.07
Sixth	3	1.33	.58	4.67	.58	3.67	1.15	1.33	.58	2.33	1.53
Seventh	1	3.00	00	4.00	.00	4.00	.00	2.00	.00	2.00	.00

[a] numbers based on a Likert-scale with 1=Never to 5=Always.

Appendix I[1]

AAR Time Data

Table I-1

Total AAR Time

Training Table	n	M	SD
First	8	31.75	4.59
Second	9	31.44	10.89
Third	7	26.00	6.78
Fourth	6	26.17	8.06
Fifth	7	23.14	10.79
Sixth	3	27.00	7.00
Seventh	1	19.00	.00

Table I-2

Proportion of AAR Time for O/C Lecture

Training Table	n	M	SD
First	8	.37	.23
Second	9	.42	.18
Third	7	.30	.13
Fourth	6	.28	.09
Fifth	7	.23	.16
Sixth	3	.27	.13
Seventh	1	.32	.00

[1] Numbers of units differed because of incomplete data.

Table I-3

Proportion of AAR Time for Participant's Lecture

Training Table	n	M	SD
First	8	.07	.07
Second	9	.07	.05
Third	7	.10	.09
Fourth	6	.07	.13
Fifth	7	.10	.09
Sixth	3	.15	.13
Seventh	1	.00	.00

Table I-4

Proportion of AAR Time Spent for Group Discussion

Training Table	n	M	SD
First	8	.46	.24
Second	9	.43	.25
Third	7	.59	.16
Fourth	6	.64	.15
Fifth	7	.64	.14
Sixth	3	.54	.17
Seventh	1	.68	.00

Table I-5

Proportion of AAR Time Spent on Discussing Background Materials

Training Table	n	M	SD
First	8	.13	.09
Second	9	.09	.06
Third	7	.08	.07
Fourth	6	.08	.06

Table I-5 CONT'D

Training Table	n	M	SD
Fifth	7	.04	.03
Sixth	3	.08	.07
Seventh	1	.11	.00

Table I-6

Proportion of AAR Time Spent on Discussing Events

Training Table	n	M	SD
First	8	.51	.13
Second	9	.47	.14
Third	7	.53	.19
Fourth	6	.61	.09
Fifth	7	.57	.26
Sixth	3	.55	.09
Seventh	1	.53	.00

Table I-7

Proportion of AAR Time Spent on Reasons for Actions

Training Table	n	M	SD
First	8	.29	.20
Second	9	.40	.18
Third	7	.33	.34
Fourth	6	.18	.22
Fifth	7	.19	.21
Sixth	3	.29	.10
Seventh	1	.53	.00

Table I-8

Proportion of AAR Time Spent on Unit Improvement

Training Table	n	M	SD
First	8	.32	.24
Second	9	.22	.10
Third	7	.22	.08
Fourth	6	.28	.19
Fifth	7	.34	.24
Sixth	3	.24	.12
Seventh	1	.37	.00

Table I-9

Proportion of AAR Time Spent on SIMNET Problems

Training Table	n	M	SD
First	8	.01	.01
Second	9	.02	.03
Third	7	.00	.01
Fourth	6	.00	.00
Fifth	7	.01	.02
Sixth	3	.00	.00
Seventh	1	.00	.00

Table I-10

Proportion of AAR Time Spent in Setting Up Feedback Materials

Training Table	n	M	SD
First	8	.02	.03
Second	9	.08	.11
Third	7	.02	.03
Fourth	6	.01	.02
Fifth	7	.04	.05
Sixth	3	.03	.04
Seventh	1	.00	.00

Table I-11

Proportion of AAR Time Spent Using Stealth

Training Table	n	M	SD
First	8	.42	.11
Second	9	.36	.14
Third	7	.46	.21
Fourth	6	.25	.23
Fifth	7	.48	.27
Sixth	3	.46	.05
Seventh	1	.63	.00

Table I-12

Proportion of AAR Time Spent Using Maps

Training Table	n	M	SD
First	8	.07	.05
Second	9	.08	.07
Third	7	.04	.04
Fourth	6	.08	.10
Fifth	7	.05	.08
Sixth	3	.08	.14
Seventh	1	.05	.00

Table I-13

Proportion of AAR Time Spent Using Posters

Training Table	n	M	SD
First	8	.06	.05
Second	9	.06	.07
Third	7	.03	.03
Fourth	6	.07	.07
Fifth	7	.03	.05

Table I-13 CONT'D

Training Table	n	M	SD
Sixth	3	.05	.06
Seventh	1	.16	.00

Table I-14

Proportion of AAR Time Spent Using Writing Board

Training Table	n	M	SD
First	8	.04	.09
Second	9	.02	.04
Third	7	.03	.08
Fourth	6	.03	.07
Fifth	7	.01	.02
Sixth	3	.04	.07
Seventh	1	.00	.00

Table I-15

Proportion of AAR Time Using Other Feedback Materials

Training Table	n	M	SD
First	8	.07	.11
Second	9	.10	.14
Third	7	.19	.24
Fourth	6	.04	.04
Fifth	7	.13	.23
Sixth	3	.00	.00
Seventh	1	.16	.00

Appendix J[1]

The Observers' Ratings for the AARs

I. SUMMARY OF QUESTIONS DEALING WITH O/Cs' COMMENTS

Table J-1

Lectured to the Group

Training Table	n	M	SD
First	8	2.88	1.13
Second	9	3.00	.71
Third	6	3.17	.98
Fourth	6	3.33	.82
Fifth	7	3.29	.49
Sixth	3	3.33	1.15
Seventh	1	3.00	.00

Table J-2

Acted as Facilitator of Discussion

Training Table	n	M	SD
First	8	3.00	.53
Second	9	2.89	.78
Third	7	3.29	.95
Fourth	6	3.33	.82
Fifth	7	3.00	.82
Sixth	3	3.00	.00
Seventh	1	2.00	.00

[1] Numbers of units differed across measures and tables because of incomplete data. Also means were based on a Likert-scale with 5=Always to 1=Never.

Table J-3

Stated Reasons for Unit's Actions

Training Table	n	M	SD
First	8	2.13	.83
Second	9	2.78	.83
Third	7	2.00	.82
Fourth	6	3.33	.82
Fifth	7	2.71	1.11
Sixth	3	2.67	.58
Seventh	1	3.00	.00

Table J-4

Stated Possible Alternative Actions

Training Table	n	M	SD
First	8	2.13	.99
Second	9	2.67	1.12
Third	7	2.43	1.13
Fourth	6	2.83	.98
Fifth	7	2.57	1.13
Sixth	3	2.67	1.15
Seventh	1	3.00	.00

Table J-5

Stated Lessons Learned

Training Table	n	M	SD
First	8	2.50	1.07
Second	9	2.67	.87
Third	7	2.43	1.13
Fourth	5	2.40	1.14
Fifth	7	2.29	1.11

Table J-5 CONT'D

Training Table	n	M	SD
Sixth	3	3.00	1.73
Seventh	1	4.00	.00

Table J-6

Discussed Skills to be Sustained

Training Table	n	M	SD
First	7	3.14	1.22
Second	8	2.88	1.36
Third	6	3.00	1.67
Fourth	6	3.67	1.03
Fifth	7	4.14	1.07
Sixth	3	2.33	1.15
Seventh	1	2.00	.00

Table J-7

Made Critical Comments about Unit's Performance

Training Table	n	M	SD
First	8	2.50	.76
Second	9	2.00	.70
Third	7	2.00	.58
Fourth	6	1.83	1.17
Fifth	7	1.57	.53
Sixth	3	1.67	.58
Seventh	1	2.00	.00

Table J-8

Posed Questions to the Entire Group

Training Table	n	M	SD
First	7	2.71	1.49
Second	9	3.55	1.13
Third	7	4.57	.53
Fourth	6	4.17	.75
Fifth	7	3.86	1.07
Sixth	3	3.67	.58
Seventh	1	4.00	.00

Table J-9

Lost Time Using Feedback Materials

Training Table	n	M	SD
First	8	1.38	.74
Second	9	2.00	.71
Third	7	1.43	.53
Fourth	5	1.40	.55
Fifth	6	1.50	.55
Sixth	3	1.67	.58
Seventh	1	1.00	.00

II. SUMMARY OF PARTICIPANTS' COMMENTS

Table J-10

Stated Reasons for Their Actions

Training Table	n	M	SD
First	8	3.13	.84
Second	8	3.38	.74
Third	7	3.29	1.25

Table J-10 CONT'D

Training Table	n	M	SD
Fourth	6	3.00	1.10
Fifth	7	3.43	.53
Sixth	3	3.00	.00
Seventh	1	3.00	.00

Table J-11

Discussed Possible Alternative Actions

Training Table	n	M	SD
First	8	2.13	.83
Second	8	2.38	.52
Third	7	2.14	.38
Fourth	6	3.17	.75
Fifth	7	2.14	.69
Sixth	3	2.33	.58
Seventh	1	2.00	.00

Table J-12

Stated lessons Learned

Training Table	n	M	SD
First	8	2.38	.92
Second	8	2.50	1.20
Third	7	2.86	1.22
Fourth	6	2.67	1.03
Fifth	7	2.43	1.13
Sixth	3	2.67	1.15
Seventh	1	3.00	.00

Table J-13

Stated Skills to be Sustained

Training Table	n	M	SD
First	8	2.00	.76
Second	8	2.38	1.30
Third	7	2.43	1.51
Fourth	6	2.17	1.17
Fifth	7	2.57	1.13
Sixth	3	2.67	1.53
Seventh	1	3.00	.00

Table J-14

Made Critical Comments about the Unit's Performance

Training Table	n	M	SD
First	8	2.00	.76
Second	8	2.00	.53
Third	7	1.57	.53
Fourth	6	1.33	.52
Fifth	7	1.43	.53
Sixth	3	2.00	1.00
Seventh	1	2.00	.00

Table J-15

Comments Prompted by the O/C

Training Table	n	M	SD
First	8	2.13	.99
Second	8	2.88	1.25
Third	7	2.00	1.15
Fourth	6	2.67	.82
Fifth	7	1.71	.95

Table J-15 CONT'D

Comments Prompted by the O/C

Training Table	n	M	SD
Sixth	3	2.33	1.15
Seventh	1	4.00	.00

Table J-16

Comments clarified by the O/C

Training Table	n	M	SD
First	4	1.75	.96
Second	5	2.00	.71
Third	4	2.00	.00
Fourth	3	3.00	1.00
Fifth	3	2.33	.58
Sixth	3	2.00	.00
Seventh	1	1.00	.00

Table J-17

Made comments about SIMNET

Training Table	n	M	SD
First	8	1.50	1.07
Second	8	1.50	.53
Third	7	1.29	.49
Fourth	6	1.17	.41
Fifth	7	1.14	.38
Sixth	3	1.00	.00
Seventh	1	1.00	.00

Table J-18

Unit Members Other than Leader(s) Participating in the AAR

Training Table	n	M	SD
First	8	2.88	1.36
Second	8	3.38	.74
Third	7	4.29	.76
Fourth	6	4.17	.98
Fifth	7	3.71	.95
Sixth	3	3.33	.58
Seventh	1	4.00	.00

Table J-19

Participants Using Feedback Materials

Training Table	n	M	SD
First	8	1.63	.74
Second	8	2.25	1.28
Third	7	1.57	.53
Fourth	6	2.00	.63
Fifth	7	1.86	.69
Sixth	3	1.67	.58
Seventh	1	2.00	.00

Appendix K

Examples of "First' and "Later" Subtasks by Successive Training Tables

Listed below are examples of specific "first" and the "later" subtasks associated with the different training tables and unit types. "First subtasks" have been operationally defined as the unit's first performance of a subtask; while, "later" subtasks" are the unit's later performance of the same subtask.

Please note that different units--within a unit type--may have performed a cited "first" or "later" subtask on different training tables. Hence, the same subtask has been listed as both a "first" and "later" subtask for the armored companies' fifth training table (see Tables K-5 and K-6).

Table K-1

Examples of "First Subtasks" by Successive Training Tables for the Armor Platoons (PLTs).

Training Table	Description of the Subtasks
First	Plt leader sends a contact report to the commander.
Second	Plt reaches SP at the specified time
Third	Plt leader inspects each vehicle's position.
Fourth	Plt leader orders the platoon to assault.
Fifth	Plt consolidates and re-organizes.
Sixth	Plt leader issues a preparatory fire commander.

Table K-2

Examples of "Later Subtasks" by Successive Training Tables for Armor Plts

Training Table	Description of the Subtasks
Second	Plt executes the action drill.
Third	Plt conducts tactical movements per OPORD
Fourth	Plt leader directs a platoon battle drill.
Fifth	Plt leader sends a contact report to the commander.
Sixth	Plt executes the contact drill.

Table K-3

Examples of "First Subtasks" by Successive Training Tables for Scout/Mech Plts

Training Table	Description of the Subtasks
First	Plt deploys in appropriate formation and techniques
Second	Plt collects information about the route.
Third	Plt provides reaction time and maneuver space.
Fourth	Plt develops the situation rapidly.
Fifth	Plt moves out of impact area.
Sixth	Plt assists maneuver elements.

Table K-4

Examples of "Later Subtasks" by Successive Training Tables for Scout/Mech Plts

Training Table	Description of the Subtasks
Second	Plt moves on designated axis/routes.
Third	Plt collects information about the route.
Fourth	Plt reports all information rapidly and accurately.
Fifth	Plt develops the situation rapidly.
Sixth	Plt maintains maximum reconnaissance forward.

Table K-5

Examples of "First Subtasks" by Successive Training Tables for the Armored Companies

Training Table	Description of the Subtasks
First	Commander controls company movement.
Second	A plt locates enemy units within their area of operation.
Third	Commander estimates the situation.
Fourth	Appropriate road march techniques are employed.
Fifth	Company executes fire support in the offense
Sixth	Company maintains enemy contact and continues the mission

Table K-6

Examples of "Later Subtasks" by Successive Training Tables for the Armored Companies

Training Table	Description of the Subtasks
Second	Company uses formation that maintain security.
Third	Company conducts actions on contact.
Fourth	Commander controls company movement.
Fifth	Company executes fire support in the offense
Sixth	Commander controls company movement.

Appendix L

Data Analyses for Assessment B

Wilcoxon signed-rank tests for matched pairs were computed based on the O/Cs' first and last ratings (see Table 7) for subtasks performed at least twice. These tests were done for all units combined, and separately for armor platoons, mech/scout platoons, and armor companies. The first set of tests compared the number of subtasks first rated "improve" that changed positively (last rated "sustain") relative to the number of subtasks that were unchanged (last rated "improve.") The second set of tests compared the number of subtasks first rated "sustain" that changed negatively (last rated "improve") relative to the number of subtasks that were unchanged (last rated "sustain.") The third group of tests compared the number of subtasks with positive changes to the number of subtasks with negative changes to determine if the net change was positive or negative. The latter tests also indicated if the change (gain or loss) from first to last subtask ratings was significant. All three sets of comparisons test the null hypothesis that the proportions in two categories are equal ($p = .5$) for the sampled populations of units.

With two exceptions, the tests presented in Table L-1 resulted in statistically significant differences for all units combined, and for the three types of units. In the first set of comparisons, the results indicated that a majority of subtasks first rated "improve" did change positively to a "sustain" rating. However, this was not the case for armor companies. In the second set, the results showed, without exception, that a majority of subtasks first rated "sustain" remained unchanged, with a minority of subtasks changing negatively to an "improve" rating. The third set of comparisons showed that the positive changes exceeded the negative changes, so that there was a net positive gain from first to second ratings. Again, the exception was the armor companies that had a nonsignificant test result.

Kruskal-Wallis rank tests were performed to compare the results among the groups of units. Difference scores were computed from the number of subtasks (a) with positive changes versus unchanged for those first rated "improve", (b) negative changes versus unchanged for those first rated "sustain", and (c) positive versus negative changes. The Kruskal-Wallis tests in Table L-2 were based on ranks corresponding to the order of the difference scores, including a correction for tied ranks. None of the tests were found to be statistically significant. These results require that judgement be suspended about the possible differences among the types of units. While two of the Wilcoxon tests failed to be significant for the armor companies, suggesting that the gain from training was somewhat less than that for the platoons, the present evidence is insufficient to establish that the gain was actually smaller.

Table L-1

Wilcoxon Signed-Rank Tests of Numbers of Subtasks in Categories Based on First and Last O/C Ratings

Type of Unit	n [a]	N [b]	T	Z
Positive change compared to unchanged (First rating "train to improve")				
All units	38	31	115.50	2.60***
Armor platoons	17	14	15.00	2.35**
Mech/scout platoons	11	8	6.00	1.68*
Armor companies	10	9	19.50	.35
Negative change compared to unchanged (First rating "train to sustain")				
All units	38	34	8.00	4.95***
Armor platoons	17	16	0.00	3.52***
Mech/scout platoons	11	8	2.00	2.24**
Armor companies	10	10	1.00	2.70***
Positive change compared to negative change				
All units	38	32	60.00	3.81***
Armor platoons	17	15	8.00	2.95***
Mech/scout platoons	11	9	7.00	1.83*
Armor companies	10	8	7.00	1.54

[a] Number of units. [b] Number of non-zero differences.
*$p < .10$. **$p < .05$. ***$p < .01$.

Table L-2

Kruskal-Wallis Rank Tests of Differences Among Types of Units Based on First and Last Subtask Rating Differences Scores

Difference score	n [a]	df	H
Positive change versus unchanged (First rating "improve")	38	2	1.15
Negative change versus unchanged (First rating "sustain")	38	2	2.25
Positive versus negative change	38	2	2.09

[a] Number of units.

The counts of subtasks rated at successive tables are shown in Table L-3. These data were used to compute the percents of subtasks rated "train to sustain," as shown in Figures 2-4. For both the first and later subtasks, the number of subtasks counted and units who participated decreased markedly after the fourth table. This makes the percentages for tables 5-7 relatively unreliable, especially for the small number of scout/mech platoons. Therefore, apparent trends and differences for these data must be interpreted with caution.

Table L-3

Number of First and Later Subtasks Rated at Each Training Tables

Type of subtask and rating	Training Table						
	1st	2nd	3rd	4th	5th	6th	7th
First subtasks							
Armor platoons							
Train to improve	67	21	22	14	4	2	0
Train to sustain	79	38	23	23	12	3	1
Number of units	17	17	16	13	7	3	1
Mech/scout platoons							
Train to improve	26	16	4	2	0	1	0
Train to sustain	36	27	10	2	4	2	0
Number of units	11	10	7	3	1	1	0
Armor companies							
Train to improve	32	21	7	5	8	8	0
Train to sustain	51	29	17	13	5	12	2
Number of units	10	10	10	10	9	7	1
Later subtasks							
Armor platoons							
Train to improve		23	24	11	17	12	1
Train to sustain		30	60	63	42	31	5
Number of units		17	17	16	9	8	2
Mech/scout platoons							
Train to improve		8	5	5	3	2	1
Train to sustain		15	22	35	15	13	17
Number of units		10	8	8	4	2	2
Armor companies							
Train to improve		21	17	15	23	19	2
Train to sustain		33	49	40	37	18	18
Number of units		10	10	10	10	8	2

The results of the rank tests for first tasks show no significant increase or decrease in percentages across tables for all units combined, nor in separate tests for the three types of units (Table L-4). These findings provide no evidence for an increase in subtask difficulty across tables, producing an expected decrease in the percent of tasks rated "train to sustain" when the subtasks are performed for the first time. Either there was no such increase in difficulty, or the added difficulty was offset by general transfer of training from the learning on different subtasks trained in prior tables.

For the later tasks, the trend test results in Table L-4 also do not provide strong evidence for a monotonic increase in percentages across tables. Such an increase was expected to reflect learning as the number of task repetitions increased, on average, across the tables. However, a significant increase was obtained only for the scout/mech platoons who had performed few subtasks for their later tables. Thus, even this result is possibly unreliable.

Table L-4

Trend Tests Using Spearman Rank-Order Correlations Between the Percent of Subtasks Rated "Sustain" and Training Tables

Type of Unit	R	df	t
First ratings for subtasks			
All units	.486	4	1.11
Armor platoons	.657	4	1.74
Mech/scout platoons	.429	4	.94
Armor companies	-.029	4	-.05
Later ratings for subtasks			
All units	.486	4	1.11
Armor platoons	.543	4	1.29
Mech/scout platoons	.829	4	2.96**
Armor companies	.200	4	.40

**p < .05

Using matched pairs of values for Tables 2-6 only, Wilcoxon signed-rank tests were performed to determine if the later task percentages exceeded the first task percentages, without regard to trends across tables. For all units combined, this test was significant (see Table L-5), but the results were not significant for any of the unit types tested separately. The combined result was consistent with the increase previously noted from first to last subtasks. However, these tests included the additional subtasks that were between the first and last subtasks, while omitting subtasks that were rated on the first and seventh (or

subsequent) tables. These differences in the included subtasks, and the small number of paired values available for the test, helped to account for the insignificant tests for unit types.

Table L-5

Wilcoxon Signed-Rank Tests of Differences in Percent of First and Later Subtasks Rated "Train to Sustain"

Type of Unit	n[a]	N[b]	T	Z
All units	5	5	1.00	1.75*
Armor platoons	5	5	3.00	1.21
Mech/scout platoons	5	4	3.00	.73
Armor companies	5	5	4.00	.94

[a]Number of RCVTP tables. [b]Number of non-zero differences.
*p < .10.

Appendix M

Results of the Pertinent Questionnaire Items

Table M-1

Means and Standard Deviations of the Post-Exercise Data across Tables

Items	n	M[a]	SD
Proficiency Before Training	271	3.96	1.31
Proficiency After Training	269	5.44	.83
More Proficiency after RCVTP training than previous SIMNET training	131	5.66	1.19
Improvement Due to Time in Simulators	241	5.43	1.17
Improvement Due to AARs	241	5.54	1.21
Tables became progressively more difficult	236	5.55	1.27
Our Platoon AAR Comments helped learning	212	5.75	1.15
Our Company AAR Comments helped learning	166	5.70	1.12

[a] Means for first three items based on a Likert-scale with 7 = extremely proficient to 1 = not at all proficient. Means for remaining items based on Likert-scale with 7=strongly agree and 1=strongly disagree.

Appendix N

Data Analyses for Assessment C

A 2 x 3 repeated measures ANOVA was initially computed for the proficiency estimate data. These factors were the two levels of training (pre-training versus post-training estimates) and three types of units (armor platoons, mechanized/scout platoons, and armor companies). This analysis revealed a significant interaction effect between the training and unit type factors, $F(2, 230) = 8.41$, $p < .001$. Because of the large inequality of ns among the different units, a multivariate Box M test of homogeneity of variance then was conducted. This analysis determined that Box M = 14.68 with $\chi^2(6) = 14.34$, $p = .026$, indicating significant heterogeneity of variance. Since the overall F-test might have been biased with both unequal ns and unequal within-group variances, the interaction effect was examined by the alternate method described below.

T-tests for paired samples were computed to determine the existence of any significant differences in the participants' estimates of their before-training and after-training proficiency levels. Possible "interactions" among unit-types, which were also suggested by Assessment B's data, were examined through t-tests for independent samples that adjust for unequal variance. These were computed on a difference score measure (after-training estimate minus the before-training estimates). Also, data for the unit crews were eliminated from these analyses because they were only sampled from the companies.

In Table N-1, tests of the increase in mean proficiency estimates from before to after training were significant for all units combined, and for the three unit types separately. These results demonstrate that training participants in all types of units considered the RCVTP training to have improved their tactical proficiency.

Table N-1

Comparison of Proficiency Estimates Before and After Training

Type of units	n[a]	df	SE[b]	t
All units	269	268	.08	19.55**
Armor platoons	31	30	.17	12.88**
Mech/scout platoons	35	34	.24	7.44**
Armor companies	167	168	.08	17.44**

[a]Number of participants. [b]The standard error for the difference between the two means.
**p < .05

In Table N-2, analyses of the amount of before-after change scores were significantly smaller for the armor companies compared to both types of platoons. These results paralleled those of Assessment B which suggested that armor companies benefited less from the RCVTP training than platoons.

Table N-2

Comparisons among Units regarding Before-After Change Scores

Type of units compared	df	SE[a]	t
Armor Platoons vs. Mech/Scout Platoons	64	.17 .24	1.20
Armor Platoons vs. Armor Companies	196	.17 .08	4.05**
Mech/Scout Platoons vs. Armor Companies	200	.24 .08	2.15**

[a]The standard errors of the mean for the compared units.
**p < .05.