EFFECTIVENESS OF STRUCTURED TRAINING IN SIMULATION NETWORKING (SIMNET)

David W. Bessemer and Theodore M. Shlechter
U.S. Army Research Institute, Armored Forces Research Unit
Fort Knox, Kentucky

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

ABSTRACT

With a smaller U.S. Army active force, the readiness of Army National Guard and Reserve units has greater importance. To support needed training, Congress funded a Virtual Training Program (VTP) for Reserve Component units using simulator facilities available at Fort Knox, KY, including the Simulation Networking (SIMNET) system. This research examined the effectiveness of the VTP during initial developmental trials with SIMNET exercises to validate the training approach adopted in the program. The program design established a structured framework for training across a sequence of exercises and for the training process within exercises.

The SIMNET VTP includes about 100 exercises (called tables) that provide practice on tasks critical to performing fundamental tactical operations, and to offensive and defensive missions. Subgroups of tables deliver intensive training for specific types of platoon, company, or battalion-sized units in a two-day weekend drill period. The training is guided by highly trained observer/controllers (O/Cs) who conduct the SIMNET exercises and lead after-action reviews with the participants. During table execution, the O/Cs follow detailed event guides to ensure that conditions requiring performance of specific planned tasks occur during each table exercise.

More than 75% of the National Guard and Reserve units in the VTP trials completed 4-6 SIMNET tables in a two-day training period. Measures obtained from trained observers, VTP O/Cs, and VTP participants were used as training effectiveness indicators. Observers recorded the: (a) time taken to complete a table, (b) unit tactical errors, and (c) coaching provided by the O/Cs. The O/Cs rated unit performance in each table, identifying subtasks that the units performed adequately, and subtasks needing improvement. Soldiers serving in leadership positions with the participating units estimated their unit's proficiency before and after training on a seven-point scale. Results from all of these indicators provided convergent evidence for training effectiveness.

ABOUT THE AUTHORS

David Bessemer and Theodore Shlechter are research psychologists with the U.S. Army Research Institute, Armored Forces Research Unit, ATTN: PERI-IK, Fort Knox, KY, 40121-5620. Drs. Bessemer and Shlechter are members of the Unit Training Strategies Team, currently involved in research on collective training and feedback in simulators. Paul Nesselroade and James Anthony are graduate students attending the University of Louisville. Mr. Nesselroade was a 1994 Consortium Fellow, and Mr. Anthony is currently a Consortium Fellow at the Armored Forces Research Unit. Inquiries should be directed to Dr. Bessemer (phone: 502-624-4932, e-mail: bessemer@alexandria-emh2.army.mil.).

EFFECTIVENESS OF STRUCTURED TRAINING IN SIMULATION NETWORKING (SIMNET)

David W. Bessemer and Theodore M. Shlechter
U.S. Army Research Institute, Armored Forces Research Unit
Fort Knox, Kentucky

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

INTRODUCTION

With a smaller U.S. Army active force, the readiness of Army National Guard (ARNG) and Reserve units has greater importance. These units have limited time and other resources for premobilization training. They need effective training delivered in an efficient manner. To support needed training, Congress provided funds to develop a Virtual Training Program (VTP) for Reserve Component units using simulator facilities available at Fort Knox, KY. The intent of the program has been to give these units an intensive Combat Training Center-like training experience in virtual simulation during week-end drills. The purpose of this research was to assess the training effectiveness of VTP Simulator Networking (SIMNET) exercises during the initial developmental trials using the program.

In keeping with the Congressional mandate and concepts advanced by Brown (1991), the VTP focused first on using SIMNET technology to train ARNG armor and mechanized units. The developed program has proved equally useful for Army Reserve and active units. To create the VTP, the U.S. Army Research Institute's Armored Forces Research Unit at Fort Knox managed and monitored the efforts of a contractor team to develop and demonstrate an innovative training system. This system was designed to capitalize on SIMNET capabilities, while providing necessary preparation and performance feedback components. In support, the U.S. Army Armor School procured simulation hardware and software, and trained a team of observer/controllers (O/Cs) to implement the VTP.

Trials of the prototype training exercises were conducted with ARNG and other units in the first half of 1994. Units reacted favorably to the general VTP training system. Trial findings then guided detailed revisions to the VTP exercises. The final refined program materials became available for use at SIMNET sites in November 1994. The VTP is now in full operation at Fort Knox, with O/Cs regularly training ARNG and other units. Portions of the VTP have been exported to support ARNG training at Gowen Field, ID, and Macon, GA.

The Virtual Training Program

Requirement for structured training. Army force reductions require ARNG and Reserve units to achieve higher readiness levels than those demanded in the past. However, these units face constraints making readiness goals difficult to meet. Units have only 39 training days per year, reduced by time used for training preparation, travel to training sites, and administrative activities. Units have limited access to equipment, terrain, training devices, and finances. Opportunities for live or simulated unit maneuver exercises have been too infrequent to meet the need. The VTP goal is to provide more efficient and effective training with existing and new technology.

Units have not always made best use of the available opportunities. Some units have used time in the SIMNET facility inefficiently. These units used SIMNET training hours for planning and preparation while the simulators and many soldiers stood idle. Some unit exercises have included slack periods with few events to instigate task practice. Some ARNG or Reserve units have thus completed only two to three exercises in a SIMNET weekend.

Trainers unfamiliar with the technology also have not always maximized training effectiveness. Some SIMNET exercises have been planned without specifying the tasks to be trained, or how conditions will be created to assure that tasks are performed. Doctrinally critical tasks for mission performance may have been entirely omitted from training, leaving serious gaps in unit fighting capability.
Some trainers in SIMNET have relied too much on direct exercise experience to produce the desired improvement. During exercises, restricted fields of view and fleeting events have denied participants full knowledge of what happened and limited their understanding of the dynamics of battle. The after-

action review (AAR) has proved crucial for effective collective training. The AAR has given the unit time to clarify what happened, to analyze why it happened, and to identify better procedures and tactics needed to improve future unit performance. Yet some trainers, when pressed for time, have cut short or even omitted this phase of training.

Since SIMNET was initially developed as a technology demonstration rather than as a training system, limited support was provided for training. Until recently, only one control station had a 3-D view ("Stealth" display) of the battlefield enabling trainers to fully monitor and control an exercise. Also, only two classrooms with replay capabilities were available for AARs. When several units were training in SIMNET, some trainers were forced to control exercises from workstations that only had map displays. Then they led AARs in the aisles between simulators using easel pads.

Structure of the VTP training tables. The VTP materials include a complete set of training support packages (training library) for about 100 structured SIMNET exercises and exercise segments (called tables). The library contains exercises for armor battalions and battalion task forces, and tables for armor companies, company teams, cavalry troops, armor platoons, mechanized infantry platoons, and scout platoons. Also included are battalion staff exercises in both JANUS and SIMNET versions, with the SIMNET version using automated capabilities to drive staff actions.

Battalion-level offense, defense, and staff exercises are based on typical Movement to Contact and Defense in Sector missions at the National Training Center (Hoffman, Graves, Koger, Flynn, & Sever 1994). These missions segment into phases with the tables based on company and platoon mission phases. Thus, the tables and exercises are vertically integrated, with companies and platoons doing the same missions and tasks on the same terrain as done by subordinate units in battalion-level exercises.

Six table types focus on the specific tasks critical to performing fundamental tactical operations, three offensive missions, and two defensive missions. These six missions have a natural order of difficulty based on the tasks performed and conditions set by factors of METT-T: mission, enemy, troops, terrain, and time available. All tasks are judged at least partially trainable in SIMNET (Burnside, 1990).

For platoons and companies, three table variants of the six missions also increase in difficulty based on a crawl-walk-run concept. The VTP defined crawl-walk-run by O/C intervention, and exercise complexity. O/C intervention to prompt or coach the unit is allowed at three levels: as often as needed, occasionally, or rarely. Complexity is varied by the number of tasks and the speed of execution required by the tactical situation. The resulting 6 x 3 table matrix for each kind of company or platoon allows units some flexibility in choosing types of missions and levels of difficulty to meet their needs. Across the 18 tables in the matrix, tasks are usually repeated three or more times.

Training process and procedures. The VTP is implemented in a structured framework of activities. The VTP provides "turn-key" training management to reduce the burden of planning and training delivery imposed on units. Before a unit rotates to Fort Knox, O/Cs visit to brief the unit and to provide a VTP guide, orders, and map overlays. After the initial VTP trials, O/Cs also provide videotapes illustrating successful table execution to help the unit prepare for VTP training. SIMNET time focuses primarily on mission execution and AARs. A typical unit weekend training schedule includes four hours of SIMNET familiarization, and 12-18 hours devoted to training tables. In annual training, ARNG or Reserve units use one week to train at platoon, company, and battalion levels.

Training follows a strict agenda to control time and assure efficiency. Times allowed for table preview, troop-leading, execution, and AAR activities total about two hours per table. Tables are designed to be executed in about one hour with little slack time. Battalion exercises are 4-6 hours. The VTP training is guided by experienced O/Cs (most with captain rank) who conduct previews of table tasks and terrain, control the SIMNET exercises, monitor performance, and lead the AARs. Each O/C is assisted by an NCO or civilian Exercise Controller (EC) who controls the opponent force and other simulation events. Both O/C and EC work at an O/C station equipped with radios, Plan-View Display (PVD) and Data Logger workstations, and three large monitors showing an out-the-window "Stealth" 3-D battlefield view. The Fort Knox SIMNET facility was upgraded with 12 O/C stations to support VTP exercise control and replay in AARs.

The VTP training is structured in more detail by table Event Guides, and an AAR Agenda. An Event Guide lays out O/C messages and EC actions that create initiating conditions to assure the unit opportunities to perform specific tasks during the exercise. The expected unit actions also are listed

with room for the O/C to record his comments. The AAR Agenda and the O/C's AAR instructional techniques are based on the discovery learning model (Department of the Army, 1993). The O/Cs follow up on each unit's training by preparing a take-home package (THP). The THP reviews the unit's VTP training results and summarizes lessons learned. The unit can use the THP to identify training needs and plan further training.

Lessons Learned. The VTP development and initial trials produced several lessons. First, the turn-key approach conflicted with Army tradition that places full responsibility for training on the unit commander. However, some expected objections were voiced only at battalion level, where the unit commanders desired more freedom to plan their own mission course of action and orders. Second, the vertical integration of exercises proved to be repetitious and predictable when units trained at each echelon in sequence. To retain novelty, variation in terrain and opponent actions is advisable. Third, repeated runs through tables and revisions of the Event Guides were required. To make tables work as designed, the sequence and timing of events had to be carefully adjusted. Fourth, given current technology, both an O/C and an EC were necessary to control effective exercises and conduct AARs. The workload and pace of VTP training exceeded what one O/C can manage alone.

Research Objectives

The primary objective of this research was to assess the effectiveness of the VTP training to validate the approach adopted in structuring the program's content and process. A secondary objective was to examine performance measures from several sources that might be used as training quality indicators.

Evaluation Approach

Previous evaluations of SIMNET training. Most evaluations of SIMNET training effectiveness have been inconclusive because of small sample size or other methodological problems (Boldovici & Bessemer, 1994). For example, performance gains after five days of SIMNET training were found for tank platoons and for mechanized infantry platoons (Test and Experimentation Command [TEXCOM] Combined Arms Test Center, 1990). Measures were based on go/no-go ratings of task and subtask performance in field exercises before and after SIMNET training. These gains suggested positive transfer from SIMNET training to later field performance. However, without control groups, what part of the gains resulted from pretest practice or from SIMNET training could not be determined.

Evidence of positive transfer was obtained for platoon leader training in the Armor Officer Basic (AOB) Course in the Armor School (Bessemer, 1991; Shlechter, Bessemer, & Kolosh, 1991). Using an interrupted time series design, baseline classes were compared to later classes after adding two days of SIMNET training to the AOB course. Instructor ratings of leader performance in field exercises at course-end were used to assess transfer. The results underscored the role of performance feedback and task-specific training in producing transfer effects. Bessemer found that transfer increased gradually over several months as the AOB instructors gained SIMNET training experience. His observations suggested that this change was associated with improved feedback provided to the AOB students during AARs. Shlechter, et al. showed that students trained in leadership positions (platoon leader or sergeant) in SIMNET exercises had greater transfer than those who trained only as tank crewmen.

Multiple effectiveness indicators. The purpose of the developmental trials was mainly to conduct a formative evaluation to improve the VTP procedures and tables. Training conditions in the trials could not be experimentally varied or controlled, nor could transfer be measured in field exercises. To avoid disrupting training, this research relied on observation and unobtrusive data sources for evidence to support inferences about VTP effectiveness. Several different indicators obtained by different methods were examined. These indicators provided three different perspectives on collective training outcomes. First, independent observers recorded aspects of the training that may reflect the proficiency of units. Second, the VTP O/Cs assessed the unit's performance of tasks performed in the training tables. Third, the unit leaders reported impressions about what their units may have gained from the VTP training. Inferences about the effectiveness of the VTP could be supported if similar results for different indicators converged on a common conclusion.

OBSERVATIONS OF TRAINING

Method

Unit sample. Nine units were each observed during one and one-half days of VTP training. The units included 3 armor companies, 3 armor company teams, and 3 armor platoons. Most were ARNG units, but two companies and one team were active units. Seven units completed at least five tables, but

only three completed six tables. Of the 45 tables observed, just two were defensive.

Procedure. Three Army Research Institute (ARI) research staff members observed the selected units' training. These observers were not involved in VTP development. Each observer collected data for 4-8 hours at different times for each unit to cover the VTP training, omitting the initial four hours of SIMNET familiarization. The observers recorded time and events, and rated aspects of training on prepared forms during all phases of training.

The observers watched the table preview and AAR presentations at the O/C station behind the audience formed by the unit personnel. The observers viewed each unit's performance on the displays at the O/C station and listened to radio communications on the unit nets. The time and actions of the O/C, EC, and unit during table execution were noted in relation to the table event guide, along with any SIMNET downtime that delayed the exercise.

Criterion measures. A number of measures derived from the observations were examined. Three measures assumed to be sensitive to training effects were selected as prime indicators of unit proficiency These measures were:

1. Exercise Time--minutes elapsed from the O/C's first command starting table execution to the last action completed at the end of the table, with SIMNET downtime subtracted from this measure.

2. Tactical Errors--number of errors of commission or omission in commands or actions involving movement, navigation, engagement, or reporting to the O/C. Errors in task execution were readily apparent in relation to the unit actions outlined in the Event Guide.

3. O/C Coaching--number of interventions by the O/C to direct or prompt action by the unit. Some examples of coaching were prompting proper radio procedures, simulating adjacent unit reports to signal incorrect unit position, sending artillery fire to stimulate movement, and requesting a report.

Results

Analysis. Unit variation in the number of completed tables prevented any simple analysis of measures across successive tables. Therefore, trends were measured for individual units based on the tables available, and then combined over units.

Rank-order correlations (Kendall's τ) were used to quantify trends in each units' exercise time, errors, and coaching measures across successive tables. A unit's values were ranked for each measure, and a unit's order of completing tables. Correlations were computed between table order and measure rankings for each individual unit with tied rank corrections (Kendall, 1975). Two-tailed weighted t-tests were then done to determine if the mean τ values differed from zero, indicating significant average trends.

Exercise time. As training progressed, units used less time to complete the VTP tables. Shown in Figure 1, time for the first table was more than twice that for later tables, and most of the decrease occurred by the third table. This negative trend for time found across successive tables was significant. The mean $\tau = -.574$, and $t(8) = 6.31$, $p < .001$.

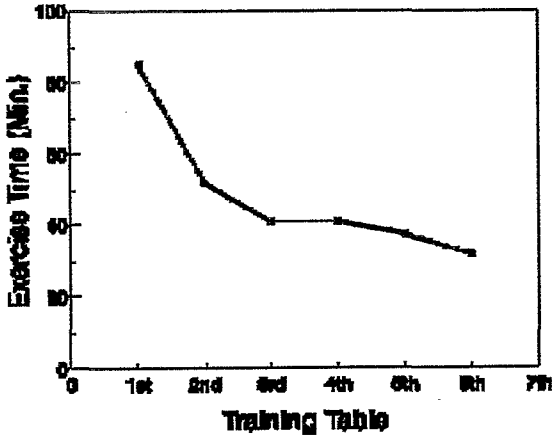

Figure 1. Mean time to complete VTP tables.

Number of tactical errors. Errors decreased in a pattern similar to that for time. Units averaged 12.9 errors on their first table, but only 5.1 for the seven units completing their fifth table. Again most of the decrease occurred by the third table. The negative trend for the error measure was significant, with $\tau = -.340$, and $t(8) = 5.86$, $p < .001$.

Observer/Controller coaching. The frequency of coaching declined from 8.4 on the first table to 3.4 on the fifth table, also mostly by the third table. The negative trend for the coaching measure was significant, with $\tau = -.206$, and $t(8) = 3.32$, $p < .02$.

Summary

Units typically took less time, made fewer errors, and needed less coaching as they completed successive VTP tables. These measures suggested that the units' performance improved as training progressed. The decreasing trends were observed even though the

designed difficulty of the tables increased. However, most of the improvement was found only up to the third table. Increased difficulty may have partly offset gains on tables after the third. On the other hand, these results alone are inconclusive, since some part of the change on the early tables might be attributed simply to units adjusting to the SIMNET equipment and visual display of terrain and targets.

OBSERVER/CONTROLLER JUDGMENTS

Method

Unit sample. Judgment data were obtained for 38 units. The units included 17 armor platoons, 10 armor companies, 5 scout platoons, and 5 mechanized infantry platoons. Most were ARNG units, but 4 armor companies and 1 scout platoon were active units. These units executed 32 fundamental, 89 offensive, and 66 defensive training tables, totaling 187 tables. The median number of tables per unit was about 5.5.

Procedure. During VTP table execution, O/Cs noted the units' performance of tasks and subtasks listed in the Event Guides. The O/Cs later used their notes to rate subtasks which units needed either to "train to sustain" or "train to improve," i.e., judging performance as satisfactory or unsatisfactory. These ratings for each table were provided to the unit in its THP. We collected copies of the THPs, and used the ratings to measure unit performance.

Criterion measures. The number of subtasks in the "sustain" and "improve" categories were counted for each unit and totaled across units. The percent of subtasks in different categories measured the unit's level of performance. Since the same subtasks did not always repeat in consecutive training tables, the numbers and percentages were done in two ways:

1. First and Last Subtask Ratings--for all subtasks that a unit performed at least twice. the numbers and percents were determined for the four possible combinations of first and last ratings. i.e., "sustain-sustain," "sustain-improve," "improve-sustain," and "improve-improve." These measures reflected the changes in performance from the start to the end of the unit's subtask training, without regard to the particular tables in which the subtasks occurred, or how many times they occurred. These measures ignored subtask ratings between the first and last.

2. First and Later Subtask Ratings--for all subtasks performed by a unit, the number of subtasks in each rating category were counted for successive training tables and totaled for each type of unit. These numbers also were determined separately for each unit's (a) first performance of a subtask (first subtasks) and (b) all later performances of the same subtask (later subtasks). Percents of subtasks rated "sustain" were then computed for successive tables both for first and later subtasks. The percents for the first subtask ratings measured the unit's subtask proficiency prior to the VTP training, and also indicated whether that proficiency was stable over successive tables. The measures based on later subtasks indicated the cumulative gain across tables from repeated subtask training.

Results

Analysis. The small samples of mechanized infantry and scout platoons were combined together in all analyses. Wilcoxen signed-rank tests (Lehmann, 1975) were used to compare unit counts between various pairs of categories for counts of first and last ratings, This tested if the units tended to deviate from equality in the same direction. Kruskal-Wallis rank tests were used to test if the frequency differences between categories were equal among types of units.

With first and later subtask ratings, Spearman rank correlations were used to test trends in performance across training tables. Analyses of variance and covariance were also performed on first and later ratings combined over tables by units. These analyses were used to examine the overall gain in performance, and differences among units.

Subtask performance changes. The percent of subtasks with "train to sustain" ratings increased from 61.8% to 78.6 %. Table 1 shows subtask ratings for each first-last rating combination. For 222 subtasks first rated "sustain" only a few ratings (31) changed negatively to "improve," while most were unchanged. The Wilcoxen test on unit frequencies indicated that this difference between categories was significant, $z = 4.95$, $p < .01$. For 137 subtasks first rated "improve," many ratings (91) changed positively to "sustain," and this difference was significant, $z = 2.60$, $p < .01$. The net effect was that positive changes significantly exceeded negative changes, producing the increase observed between the first and last subtask ratings. For those 122 subtasks with changes from first to last rating, 74.6% were positive changes, $z = 3.81$, $p < .01$.

Table 1

Ratings for First and Last Performance of Subtasks

	Subtask Performance

First/Last Rating Percent	First n	First Percent	Last n	Last
Sustain-	222	61.8		
Sustain			191	86.0
Improve			31	14.0
Improve-	137	38.2		
Sustain			91	66.4
Improve			46	33.6

Trends across tables. Subtask performance ratings for successive training tables appear in Figure 2. For the first subtask ratings, performance varied little over the tables. Increased difficulty of later tables could be expected to produce a decrease in initial performance level across tables, but such a trend was not found. Also, no major gain occurred for the first three tables like that found with the observational measures. The initial performance of subtasks could be expected to increase if a large part of the improved performance was simply derived from adjusting to SIMNET.

For later subtask ratings, the performance level was higher for most tables, but the expected increase across tables was irregular. Thus, evidence was lacking for a strong cumulative "learning curve" resulting from task repetition. It should be noted that the number of subtasks decreased as fewer units continued in training, making the points less reliable for the later tables. The Spearman rank correlations did not demonstrate significant trends for either first or later subtask performance ratings.

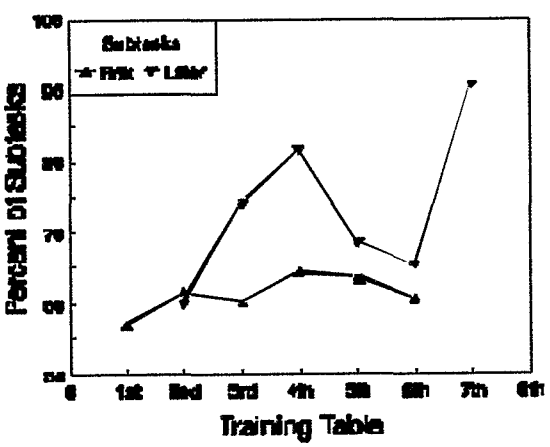

Figure 2. Performance trends measured by subtasks rated "sustain" over successive VTP tables.

Differences among unit types. For the first and last ratings, armor companies had smaller gains (60.5% to 71.1%) than the platoons (62.4 to 82.4). Also, the decrease in later subtask performance for the fifth and sixth tables in Figure 2 was larger for companies than for platoons. However, no unit type comparisons were significant, indicating that the apparent differences were unreliable.

The analysis of the first and later ratings combined over tables did not reveal any significant differences among the different types of units. Repeated measure analysis of variance showed a significant gain from first ratings (58.9%) to later ratings (70.5%), $F(1, 34) = 12.44$, $p < .001$, but no other effects. This result reconfirmed the gain found for first and last ratings. Analysis of covariance showed no differences in regression or adjusted means among unit types when the first task percents were used as a covariate with the later task percents used as the dependent variable.

Initial unit proficiency. In the covariance analysis the regression effect was significant ($r = .62$, $t(33) = 4.64$, $p < .001$. The r^2 indicates that about 38% of the variance in later subtask ratings was predictable from the first ratings for those subtasks. The regression coefficient was .788 with a 95% confidence interval ranging from .442 to 1.134. Thus, for every 10% increase in first ratings, the later ratings increased about 7.9%. This relationship demonstrated that a unit's performance in VTP training was very dependent on its proficiency level entering training.

Summary

The measures based on O/C ratings of subtask performance, whether first and last ratings or first and all later ratings, showed that substantial gains resulted from the VTP training. These measures also showed that the gains are subtask specific, and not a general effect of adjusting to conditions of the simulation. On the other hand, task repetition was not seen to produce consistently larger gains on the later tables that were completed. Furthermore, performance remained dependent on the initial proficiency of the unit. These findings implied that units can profit from more than two days of SIMNET training. Some units may need much more training to fully overcome an initially low level of subtask performance.

UNIT LEADER OPINIONS

Method

Sample. A questionnaire was administered to 233 soldiers serving in leadership positions within units trained in the VTP developmental trials. The positions included company commanders, executive officers, platoon leaders, platoon sergeants, and tank or vehicle commanders.. By unit type, 167

respondents came from 19 armor companies, 31 from 12 armor platoons, and 35 from 3 scout and 3 mechanized infantry platoons.

Questionnaire and measures. The VTP contractor team prepared a questionnaire form with some input from ARI. This form contained 64 items related to nine topics about formative evaluation issues concerning the VTP. Results are presented here from just two items that addressed unit proficiency for tasks in VTP tables completed by the unit. Only respondents who completed both items were included in the present sample. These seven-point items asked the leaders to estimate for the VTP tasks performed:

1. Unit proficiency before VTP training--with 7 labeled as "extremely proficient", 1 labeled as "not proficient" and 4 labeled as a "neutral" point. The other numerical points were unlabeled.
2. Unit proficiency after VTP training--with the same number scale and labels as the previous item.

Procedure. The questionnaire form was passed out at the start of the second day of training, and picked up at a post-training debriefing conducted by VTP contractor personnel. The respondents were asked to complete the form whenever they had time during breaks in training. Contractor personnel entered the responses in a database provided to ARI.

Results

Analysis. Analysis of variance was performed with the before-after proficiency estimates as a repeated measure factor and unit types as a between-subjects factor. Significant effects were examined further by pairwise comparisons. Paired sample t-tests were done to compare before with after mean estimates. Independent sample t-tests on before-after difference scores were done with adjustments for unequal variance to compare means for types of units. Since sample sizes differed greatly among the groups, and the Box M-test done with the analysis of variance indicated heterogeneity of variance, χ^2 (6) = 14.34, $p = .026$, the α-level for tests assuming equal variance could be seriously biased.

Estimates of proficiency level. The mean estimates of proficiency made by leaders in each type of unit are shown in Figure 3. These estimates increase from before to after in all three groups. An analysis of variance revealed a significant interaction between the before-after estimate and unit type factors, $F(2, 230) = 8.41$, $p < .001$. Differences between before and after estimates (gains) were significant for each unit type. The test results were: for the armor platoons, $t(30) = 12.88$; for the combined mechanized infantry and scout platoons, $t(34) = 7.44$; and for the armor companies, $t(166) = 17.44$, all with $p < .001$.

Differences among unit types. Comparisons of difference scores showed that the before-after difference was smaller for armor companies. The types of platoons did not differ significantly, $t(64) = 1.20$. However, the armor companies differed both from armor platoons, $t(196) = 4.05$, $p < .001$, and from the combined mechanized infantry and scout platoons, $t(200) = 2.15$, $p < .05$. These differences in estimated gains also reflected the fact that the before-training ratings were higher for companies compared to both types of platoons, whereas all groups had similar after-training ratings.

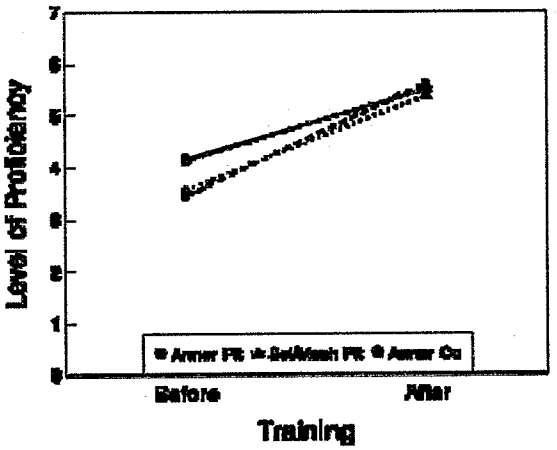

Figure 3. Leader estimates of unit proficiency level.

Summary

The unit proficiency estimates made by unit leaders, on the average, indicated their belief that their units improved as a result of VTP training. This was the result for all types of units, in agreement with other indicators examined in this research. However, the before-after estimate differences were larger for platoons than for companies. This result agreed in part with smaller company gains (but not statistically supported) found in the O/C's subtask performance ratings. If confirmed in future, this finding would imply that more company-level training is needed to match the gains found for platoon-level training. However, such effects may have reflected changes in the O/Cs' or unit leaders' judgment scales at these levels instead of actual performance differences. The O/Cs' judgments disagreed with company leaders' opinions about unit proficiency before training, since the O/Cs' initial performance ratings for subtasks were no better for companies than for platoons.

DISCUSSION

Agreement Among Effectiveness Indicators

The findings of this research point to one common outcome: completing VTP tables produces gains in unit performance. The measures obtained from observer's records, O/C's judgments, and leader's opinions showed significant changes demonstrating improved unit proficiency. Although alternative interpretations are possible, the consistency among different sources and measures encourages the inference that the VTP training caused the changes.

The results are more ambiguous about the extent and pattern of the improvement. Several explanations are possible for variations found among measures. One is simply that they measure somewhat different things. The samples and possible confounded variables differ for each source of data. Some additional controls and measures are required to disentangle the conditions that might be responsible for different results.

The indicators do agree that additional improvement in performance is possible in all cases. None of the measures approached the upper limit of their scales. Also, platoons may have improved more in VTP training than companies. These findings could reflect real effects of task difficulty or they may be traceable to other artifacts. Further evaluation of the VTP training is needed to project how much training is required to reach high degrees of proficiency.

Implications for Program Implementation

Experience gained in the initial VTP trials and the evidence of training effectiveness help to establish some conditions for successful unit training in virtual environments. The VTP demonstrates that well-designed structured training can assure both efficient and effective training. Much past experience with individual and crew trainers has confirmed this fact, and the VTP extends the principle to unit training.

First, efficient training results from deliberate design decisions that increase the productive use of training time. Providing prepared orders to reduce on-site planning, following a disciplined training agenda, and reducing slack time in table execution all contribute to the efficiency of the VTP.

Second, dedicated trainers are required for effective training delivery. Training functions might be better automated to reduce the workload and remove the EC. However, an O/C's technical and military expertise will still be needed to conduct complex tactical exercises and lead AARs. Unit personnel who infrequently conduct such training lack the needed expertise.

Third, to gain the full benefit of structured training, units must be able to train often. ARNG units, for example will take three weekend drills just to execute once all 18 VTP library tables for their type of unit. Also, the present results imply that repetition of some tables or variants with similar tasks will be necessary to attain higher performance levels.

Training Quality Indicators

The indicators used in this research all proved to be sensitive to the effects of training, and are good candidates for monitoring training quality. The use of such indicators would provide a means of assuring that the training effectiveness of the program is maintained as the original O/C cadre are replaced. Long term records would enable retention loss or cumulative gains in unit proficiency to be measured as units return repeatedly for training. The indicators could also assess program changes made in an ongoing quality improvement effort.

Both the O/C ratings and leader opinions can be obtained without interfering with the training process, and without additional cost. The observer measures require additional personnel. Although the O/Cs and ECs might be asked to collect such measures, they are not independent observers, and the additional task could interfere with performance of their training duties. Other measures derived from network data could provide useful indicators, but obtaining and processing such data requires more analytical support than is presently available.

CONCLUSIONS

The VTP structured training is more efficient than some past SIMNET training, enabling more exercises to be completed in a given time period.

The structured VTP training is effective to some degree given results in the initial trials of the program. The full effectiveness of the refined implemented program remains to be determined.

Measures of unit performance that are sensitive to training effects are available for monitoring the quality of the implemented program.

REFERENCES

Brown, F. J. (1991). A simulation-based intensified training readiness strategy for the reserve component (IDA Paper P-2611). Alexandria, VA: Institute for Defense Analysis.

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. (DTIC No. AD A233 198)

Boldovici, J. A., & Bessemer, D. W. (1994) Training research with SIMNET: Implications for designing research with distributive interactive simulations (ARI Technical Report 1006). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (DTIC No. AD A285 584)

Burnside, B. L. (1990). Assessing the capabilities of training simulations: A method and simulation networking (SIMNET) application (ARI Research Report 1565). Alexandria, VA: U.S. Army Institute for the Behavioral and Social Sciences. (DTIC No. AD A226 354)

Department of the Army. (1993). A leaders guide to after action reviews (TC 25-20). Washington, DC: Author.

Hoffman, R. G., Graves, C. R., Koger, M. E., Flynn, M. R., & Sever, R. S. (1995). Developing the RCVTP: History and lessons learned (ARI Research Report 1675). 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.

Lehmann, E. L. (1975). Nonparametrics: statistical methods based on ranks. San Francisco, CA: Holden-Day.

Shlechter, T. M., Bessemer, D. W., & Kolosh, K. P. (1991). The effects of SIMNET role-playing on the training of prospective platoon leaders (Tech. Rep. No. 938). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (DTIC No. AD A224 913)

Test and Experimentation Command (TEXCOM) Combined Arms Test Center. (1990, August). Close combat tactical trainer (CCTT) force development testing and experimentation (FDTE) (TCATC Test Report No. FD 0200). Fort Hood, TX: Author.