AN EXPERIMENTAL DESIGN TO COLLECT HUMAN SCIENCE DATA FOR MODELING AND ANALYSIS

Eugene Dutoit and William Guest, Dismounted Battlespace Battle Lab, Fort Benning, GA
Michael Statkus, Natick Soldier Center
Arthur Garrett, Army Materiel Systems Analysis Activity
Luci Salvi, Army Research Laboratory

ABSTRACT

There is little data available to the modeling and analysis community for describing soldier performance in close combat/MOUT environments. Therefore, the objectives of this experiment were to obtain dismounted soldier performance data (e.g. target engagement and weapons firing) while learning about the process of collecting human performance data in a virtual combat environment. Because this project is on-going, this paper will describe the experimental design procedure, the data collection, and the follow-on statistical analysis procedures without presenting the actual data gathered.

INTRODUCTION

It is widely recognized by the Department of Defense and by the Department of the Army that many future battles will almost certainly unfold as close combat or military operations in urban terrain (MOUT). This prediction for close combat/MOUT engagements has been evidenced by the Army's recent MOUT Advanced Concept Technology Demonstration and the subsequent investment in MOUT technologies and equipment over the past few years. Of particular concern to the modeling community are the existing data gaps that need to be filled to more accurately describe engagements closer than 25 meters. Given the emphasis on close combat/MOUT and the need for more complete underlying data, this experiment was the first step in a multi-year effort to address this lack of data in the broad areas of move, shoot, and communicate; human behavior representation; and enabling data such as metabolic work load and fatigue. Ultimately, this data collection effort will support the modeling and analysis community's ability to conduct technology assessments, equipment trade-offs, and Basis of Issue analyses for the dismounted warrior.

This experiment was a joint effort involving the Natick Soldier Center (sponsoring agency and project lead), the Army Materiel Systems Analysis Activity, the Army Research Laboratory, and the Simulation Division of the Battlelab at Fort Benning, Georgia.

GENERAL EXPERIMENTAL OVERVIEW

The virtual experiment was conducted with 10 soldiers stationed at Fort Benning during a 2-week period of time. There were four tests/experiments scheduled with an overall total number of 282 replications. These separate tests used individual soldiers and two-man fire teams immersed in a virtual MOUT environment with the humans pitted against the computer-generated forces. A special data collection and analysis tool was developed to extract the experimental information from the computers logger files and put these data into spreadsheets suitable for analysis by commercial statistical programs (primarily SPSS(1) with some Excel assistance). The experimental plan summary for these four tests is presented in the table below.

TABLE 1
VIRTUAL ENVIRONMENT TEST PLAN SUMMARY

Test Number	Test Objective	Primary Issues	Number of Runs	Test Type	Lighting Conditions	# of enemy	# of Non-combatants
1	To measure a soldier's ability to detect and shoot a single computer generated force within a room during daylight operations.	Target Detection Target Engagement Weapons Firing	81	Individual soldier	Daylight	1 computer generated	0
2	To measure a soldier's ability to detect and shoot a single computer generated force within a room during nighttime operations.	Target Detection Target Engagement Weapons Firing	81	Individual soldier	Nighttime	1 computer generated	0
3	To measure a team of two soldiers' ability to detect and shoot 2 computer forces within a room during daylight operations.	Target Detection Target Engagement Weapons Firing	60	2man buddy team	Daylight	2 computer generated	0
4	To measure a team of two soldiers' ability to detect 2 targets during daylight operations and correctly identify the targets as either friend or foe and then successfully shoot the correct target (foe).	Target Detection Target Identification Target Engagement	60	2 man buddy team	Daylight	1 computer generated	1

Because the experimental design methodology for these four tests was *essentially* the same, this paper will focus on test 1 as shown above.

Virtual Combat Environment

The name of the virtual environment is the Squad Synthetic Environment (SSE). This is a squad-level man-in-the-loop simulation especially designed for dismounted Infantry applications to include individual tasks, fire team and squad level missions and urban scenarios in a virtual environment. A total of 13 full immersion soldier simulators and 10 desktop simulators are networked with computer generated forces. The SSE has been used to support analysis for training exercises, advanced concepts and requirements and research, development and acquisition. The SSE allows the dismounted soldier to move through the virtual battlefield, enter buildings, climb stairs, and move into standing, kneeling and prone positions. It also provides the capabilities for command, control, and communications. The user has a choice of terrain features to include; the McKenna MOUT site (including the details of the inside of the buildings), Camp LeJuene, dynamic terrain (put holes in buildings), detailed furniture inside of rooms and day and thermal imaging. Output analysis can be conducted on the number of rounds fired, casualties, number of targets shot, hit probability and distance to target information. Data are provided in spreadsheet format as well as video recordings.

Questionnaires

Before any of the experiments were conducted, the 10 soldiers were asked to fill out a Demographics, Experience and Training Questionnaire. The data obtained from this questionnaire will be used to provide correlated insights about the information collected on each soldier's performance in the SSE. The soldiers were also asked to fill out an After Action Questionnaire after they participated in each replication within the SSE. These questionnaires were attempting to gather additional insights from the soldiers such as; "which aiming technique did you use?"; "How difficult was it to detect the enemy?" Finally, each of the ten soldiers was asked to fill out an Exit Questionnaire when they completed *all* tests and replications. The soldiers were asked to give their general impression of the test experience and identify any problems they had with the SSE simulator.

Anthropometric Data

In addition to the questionnaire data cited above, basic anthropometric data were collected on each soldier. The study team thought that some of these measurements might be useful when determining target profiles for use in combat modeling.

TEST 1 PLAN

As stated above, this paper will focus on the "experimental design process" for this test. The other tests were similar and there is no need to repeat the same process three other times.

Objective

To measure a live soldier's ability to detect and engage a single computer-generated enemy within a room during daylight conditions.

Description

The live soldier will enter a virtual building and room that really exists at the McKenna MOUT site. The live soldier will then attempt to detect and engage (shoot) a single computer generated enemy. The enemy will be stationary but will shoot at the live soldier upon detection. Each of the nine soldiers (there are nine soldiers in an Infantry squad) will engage each of the three enemy scenario/positions in the room exactly three times. Therefore, this test will require a total of 81 replications i.e.

9 soldiers x 3 scenarios x 3 replications per scenario = 81 total replications.

Rules of Engagement for Soldiers within the Test

The following rules were followed during the course of each test. 1. Each soldier had to calibrate his weapon prior to each trial. 2. Each soldier was taught and expected to follow the standard scanning (search) techniques appropriate for city combat and to use the correct tactics, techniques and procedures. 3. Soldiers were taught to engage the enemy targets as soon as possible. 4. In order to keep soldiers from anticipating the order of presented targets they were instructed not to discuss their prior experience with any other squad members.

Pre-Test Conditions to Keep Constant

For this test, these conditions were applied to the enemy targets. 1. The enemy was stationary. 2. The enemy was placed in a kneeling position. 3. The enemy was placed behind a piece of furniture. 4. Target shape for the enemy was irregular. The right side of the body was shown from behind the furniture. 5. The enemy was set to fire on the live soldiers after being fired upon. 6. The light level was set for daytime conditions. 7. The enemy targets appeared in one of three locations/scenarios within the room; back left corner, center, or front right corner.

Measurements Obtained for Each Trial

These were the measures obtained for each trial. 1. Total time (in seconds) to detect, acquire and engage (shoot) the target. Time began when the live soldier crossed a predetermined point in the virtual room containing the enemy target. Time ended with the last trigger pull. 2. Number of virtual rounds fired by the live soldier. 3. Number of rounds fired to get the first hit. 4. Number of casualties for each trial (both live soldiers and virtual enemy.

TEST 1 PROTOCOL DESIGN AND EXECUTION

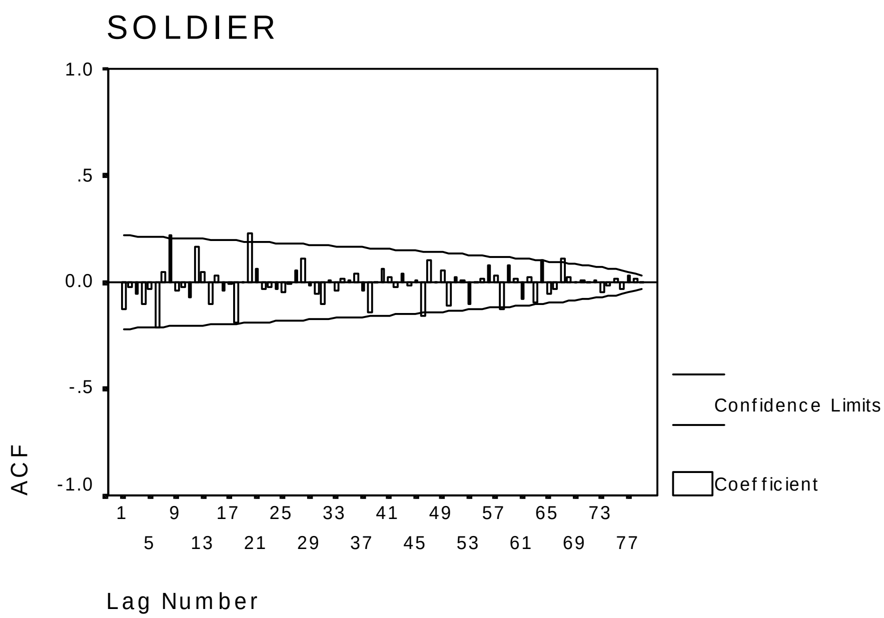
As stated above, this test was composed of three scenarios (the enemy target was located in either the left corner of the room, center of the room or the front right corner of the room). A nine-man squad of soldiers was available for the test. The number of combinations of 9 soldiers working in 3 scenarios is 27 (9 x 3). The study planners suggested that each of these 27 combinations be repeated 3 times giving a total of 81 trials for this test. A completely randomized protocol / design was constructed where each soldier was used 9 times in the test, seeing each of the 3 scenarios exactly 3 times. The randomization process was carried out using a large table of random numbers (2). For each trial (1-81) a soldier (1-9) was selected at random and then randomly assigned to a specific scenario (1-3). This process was carefully carried out until all soldiers were randomly assigned to each of the 3 scenarios exactly 3 times. The result of this randomization process is shown as Table 2 on the next page.

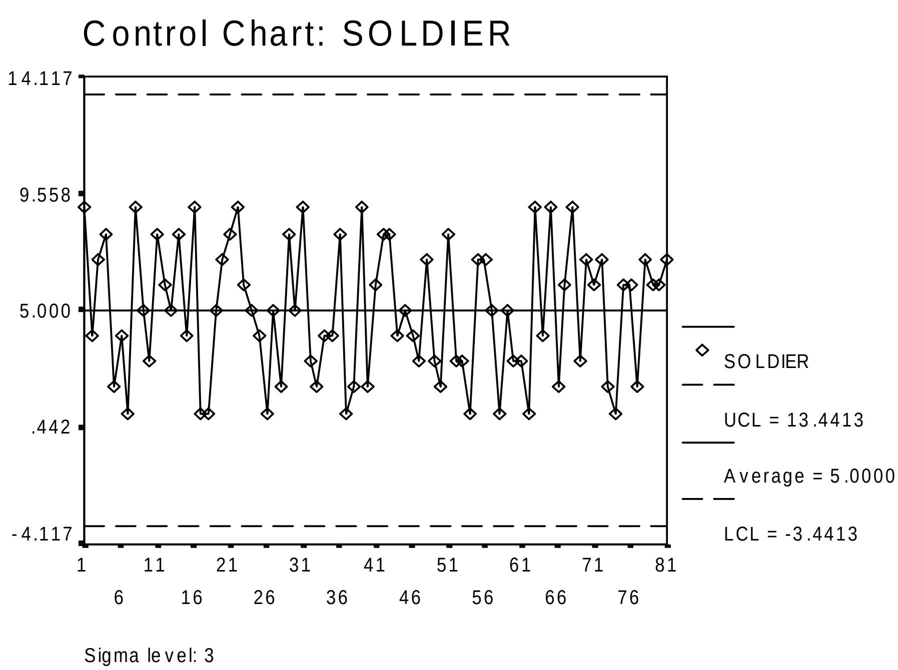
TABLE 2
RANDOM ASSIGNMENT OF SOLDIERS AND SCENARIOS TO EACH TRIAL

TRIAL	SOLDIER	SCENARIO	TRIAL	SOLDIER	SCENARIO	TRIAL	SOLDIER	SCENARIO
1	9	2	**28**	2	1	**55**	7	2
2	4	2	**29**	8	1	**56**	7	2
3	7	1	**30**	5	1	**57**	5	3
4	8	3	**31**	9	2	**58**	1	3
5	2	3	**32**	3	3	**59**	5	2
6	4	2	**33**	2	1	**60**	3	2
7	1	2	**34**	4	3	**61**	3	2
8	9	3	**35**	4	1	**62**	1	2
9	5	1	**36**	8	3	**63**	9	1
10	3	1	**37**	1	1	**64**	4	1
11	8	1	**38**	2	2	**65**	9	1
12	6	1	**39**	9	3	**66**	2	3
13	5	3	**40**	2	2	**67**	6	3
14	8	1	**41**	6	3	**68**	9	1
15	4	2	**42**	8	2	**69**	3	2
16	9	3	**43**	8	2	**70**	7	3
17	1	1	**44**	4	1	**71**	6	3
18	1	3	**45**	5	2	**72**	7	3
19	5	2	**46**	4	3	**73**	2	2
20	7	1	**47**	3	3	**74**	1	2
21	8	2	**48**	7	1	**75**	6	2
22	9	2	**49**	3	3	**76**	6	1
23	6	2	**50**	2	1	**77**	2	3
24	5	3	**51**	8	3	**78**	7	2
25	4	3	**52**	3	1	**79**	6	2
26	1	3	**53**	3	1	**80**	6	1
27	5	1	**54**	1	1	**81**	7	3

An independent verification for the randomization process was conducted. The methods used were described and recommended by the National Institute of Standards and Technology (3) by computing the autocorrealtion function for the sequences of soldiers and scenarios for the 81 trials. SPSS was used to compute these statistics. The autocorrelation functions were lagged from 1-81. There were no significant values for any of the autocorrelation functions (using a P value less than .05 as a criterion of statistical significance). In addition, a sequence quality control type plot was examined across trials (1-81) for the sequence of soldiers and scenarios to determine if there were perceptual clusters or gaps in the plotted series. None were visually apparent. For the sake of saving some space examples of these plots of the autocorrelation functions and sequence plots for the soldiers are shown in Figure 1 on the next page. The plots for the sequence of scenarios would be similar. Although the sequences of random number were drawn from a respected source, sometimes the sequence does not meet the autocorrelation criteria or perceptual quality control images. In those cases, the randomization process should be repeated and verified.

FIGURE 1
AUTOCORRELATION AND SEQUENCE PLOTS





Brief Description of the Test

This test occurred during daylight hours. One live soldier entered the virtual room. One virtual enemy soldier was positioned in the room. He was stationary and positioned randomly in one of the three locations / scenarios described in the section titled **Pre-Test Conditions to Keep Constant**. The live soldiers were randomly assigned numbers from 1 through 9 in keeping with the protocol provided in Table 2. The trials 1 through 81 were completed in the sequencing of soldiers and scenarios as described in Table 2.

A Description of Summary Statistics and Procedures

The analysis procedure was conducted using SPSS. The following steps were taken for each of the tests described in Table 1.

1. Generate box plots and histograms for the data obtained for each of the three scenarios. These graphical procedures provide information to help characterize the data; i.e. is the variable approximately normally distributed? Are there outliers or extreme values? Are the variances approximately equal? Does some of the data need to be deleted because of blunders in recording. Are nonparametric methods preferred to parametric methods? The Explore subroutine and graphics routines of SPSS are useful for doing those exploratory data analysis (4) procedures that should precede inferential data analysis.
2. List and or tabulate the summary statistics for each variable. In this experiment the statistics tabulated for the variable called “engagement time (seconds)” were; arithmetic mean, median, variance, standard deviation, minimum value, maximum value and the range of the data values.
3. If parametric methods were determined to be appropriate (data approximately normal with equal variances), then the one way analysis of variance (ANOVA) was conducted comparing the performance between the three test scenarios. The Tukey post-hoc procedure was used to isolate pairwise differences between the scenarios if the overall ANOVA indicated statistically significant differences (.05 was set as the critical value).
4. If nonparametric methods were determined to be appropriate (data not normally distributed or variances not equal between groups) then the Kruskal-Wallace test was conducted. An appropriate nonparametric post-hoc procedure for statistically significant findings using the Kruskal-Wallace test is found in (5).

A Comment Concerning Engagement Times

An initial assessment of the engagement times indicated that they were larger than expected. These expectations were based on military experience and the results of similar experiments conducted at Camp Lejeune during the MOUT Advanced Technology Demonstration (ACTD). In order to re-evaluate these engagement times each of the engagements was played back and reviewed. It was apparent that a large proportion of the total engagement time was composed of "hallway maneuver" time in the squad synthetic environment. The engagement time required from this test was supposed to consist of the time the soldier was framed (standing) in the doorway of the room plus the time he was shooting his weapon. Based on the playbacks, the "hallway maneuver" time for each trial was subtracted from each total engagement time to provide an estimate of the desired engagement time. On the average, these adjusted engagement times differed by less than 1 second from the time obtained from the MOUT ACTD live experiments at Camp Lejeune.

CONCLUSION

There is little soldier performance data available for describing close combat in an urban environment that can be used by the modeling and analysis community. The purpose of this experiment was to learn about the process of collecting these data in a virtual combat environment. This paper described the experimental design *procedure*, data collection and the follow-on statistical analysis. It was the opinion of the participants that there was excellent coordination between all the members of the squad synthetic study team. The data collection effort was successful and the numerous lessons learned will improve future efforts. This simulator study will be validated with corresponding live experiments conducted at the McKenna MOUT site.

REFERENCES

(1). SPSS Base 8.0 User's Guide, 1998.

(2). The RAND Corporation, (1955). A Million Random Digits with 100,000 Normal Deviates. The Free Press, New York, various pages.

(3). National Institute of Standards and Technology web site http://www.itl.nist.gov/div898/handbook/index2.htm .

(4). Tukey, John W. (1977). Exploratory Data Analysis. Addison-Wesley Publishing Company, Reading, various pages.

(5). Conover, W.J. (1980). Practical Nonparametric Statistics, 2 ed. John Wiley & Sons, New York, 230-231.