THE DEVELOPMENT OF COMBAT RELATED MEASURES FOR SMALL ARMS EVALUATION

RONALD D. KLEIN AND CHARLES B. THOMAS*, COLONEL, INFANTRY
UNITED STATES ARMY INFANTRY BOARD
FORT BENNING, GEORGIA

INTRODUCTION

The United States Army Infantry Board is currently conducting a 5-year Infantry Weapons Methodology Study. The aim of the methodology study is to provide test procedures and techniques which will insure the selection of the most effective weapons and equipment for the Infantry soldier. The approach of the methodology study was to cast these procedures in terms of the environment in which the candidate weapons and support equipment will be used. Since a realistic evaluation of weapon performance cannot be undertaken with validity in a sterile laboratory situation, the movement in recent years has been towards tactical or operational testing. This paper will relate some of the results thus far achieved and will emphasize in particular a new analytical technique for isolating particularly important variables or measures of effectiveness.

An objective of the methodology study is to evaluate systematically those factors which influence the combat environment by using the scientific method for as much objectivity as possible.(1) To relate the test environment to the combat environment for the rifle portion of this study, a listing of the various combat actions and tasks normally accomplished by the Infantry was prepared. As a result of this research, 26 separate combat actions were identified, such as counterambush, close combat, frontal attack, etc. Next, after researching pertinent doctrinal and training literature, a list of tasks normally accomplished by the Infantryman when executing the combat actions was prepared.(2) These include such actions as medium to short range sustained fire, rapid reaction firing and close range high intensity fire. It was determined that three basic tactical situations (attack, quickfire, and defense) would accommodate the 26 combat actions as well as the 23 combat tasks

Thirty-six measures of effectiveness were developed from a consideration of these combat actions and tasks which reflect soldier/weapon system performance under combat conditions. Examples of these are: time to first round, time between trigger pulls, distribution of near misses, time to shift fire, and hits per pound expressed as a percent of the soldier's basic load. These replace the relatively sterile measures previously used with the service test and evaluation of weapon systems, such as the stationary bull's-eye target associated with the old known-distance ranges. The new measures provide definitive descriptions related to weapon performance which assist the decision-maker in selecting the best of several competing weapon systems.

Another result of the testing thus far has been the development of a new technique for relating these specific measures to combat effectiveness. If successful, this new method will isolate the more important measures so that their impact on the decision-making task will have a weight proportional to their importance in the combat environment. The paragraphs below describe comparative weapon performance when some of the new measures are used and the technique for determining the relative value of these measures as predictors of weapon performance.

EXAMPLES OF OPERATIONAL MEASURES OF EFFECTIVENESS

The purpose of Quick-fire Experiment I was to identify and isolate factors critical to man/weapon system evaluation and to develop methods for quantitatively measuring those factors.(3) Two examples of these new measures of effectiveness are presented below. The test vehicles were two different automatic rifles. While the results appear to address the performance of these two weapon systems, the aim of this paper is to demonstrate the capability of the test facility to resolve statistical differences between weapon systems--not to evaluate the relative merits of the tested weapons used in the quick-fire experiment.

Time-To-First-Round. The time to fire the first round is defined as the time period between the instant when a target appears and the instant the first round is fired. The time-to-first-round is indicative of the actions necessary for a soldier carrying his rifle at the ready position to acquire and engage a surprise target. Time-to-first-round measures the soldier's actions of bringing the weapon to his shoulder, seating the stock against his shoulder, aligning his head so that the sights are placed in a line between the eye and the target, aligning the sights, acquiring a "sight picture," and squeezing the trigger. Should the design of the weapon inhibit any of these actions by the soldier, the effects should be reflected by this measure. The quick-fire test facility was able to find statistically significant differences between the two candidate test rifles of less than three tenths of a second for specific engagements.

Graphical presentation of time-to-first-round as a function of range is shown in Figures 1 and 2. It can be seen that Rifle A is faster than Rifle B and that both weapon systems fire more rapidly in the automatic mode than in the semiautomatic mode. Performance at specific ranges is not notably different, although, in general, the mean values increase with range. The important factor from the viewpoint of test methodology is that a specific hypothesis can be confirmed or rejected with respect to which weapon characteristics contribute to speed in firing the first round. In this test, it could be hypothesized that Rifle B, the low recoil weapon, would be the faster weapon. The test results do not confirm this hypothesis but show that the heavier weapon is approximately three tenths of a second faster, which is significant at the 5-percent level. Weapon weight and low recoil may in fact contribute to speed. Therefore, some other variable or characteristic of these two weapons systems must have an overriding influence.

It is, of course, possible to theorize as to the cause for the observed differences. For example, sighting systems may be responsible for rifle differences; Rifle B has elevated sights which may prohibit the firer from using the barrel as an aiming aid and therefore may be less suited for the quick-kill or rapid-fire technique than Rifle A. With reference to the differences between modes, there may be a characteristic psychological factor which induces more deliberation in the semiautomatic mode. The firer may feel less need to aim carefully when firing in the automatic mode since he feels that this mode offers greater fire power--enough to compensate for a hasty trigger pull. However, other measures such as hit probability show that the automatic mode is much less effective than the semiautomatic mode. This difference will be discussed later under the analysis of burst size.

Time-Between-Trigger-Pulls. The term time-between-trigger-pulls has been adopted so that reference could be made to both automatic and semiautomatic modes of fire with the same term. (Normally, the term time-between-bursts was used for automatic fire, and time-between-rounds was used for semiautomatic fire.) The term is used as an indicator of the soldier's ability to absorb the recoil, reacquire the target, obtain a new sight picture, or point the weapon and squeeze the trigger.

Figure 3 presents results of the time-between-bursts analysis. The figure shows a 25-percent difference, significant at the 5-percent level, between the two weapon systems indicating that this measure has some degree of criticality to weapon performance evaluation in the automatic mode. No significant differences were observed in the semiautomatic mode.

EFFECTIVENESS EVALUATION

The development of these measures of effectiveness has presented the decision-maker with quantitative measures of weapon performance obtained in an operational environment. This development represents a distinct advance in objective evaluation but in turn poses a new question: Which of these measures should receive the most weight in the process of choosing the optimum weapon system, especially when one candidate weapon proves superior in some measures and inferior in others? Specifically, the basic problem is to relate differences found within the various measures of effectiveness to meaningful differences in the real world. In other words, should one weapon prove to be 1.2 seconds faster with such a measure as time-to-shift-fire, how does this difference relate to the success or survivability of the soldier/weapon system in combat, and how should such a difference be equated to differences in other measures such as hit probability?

The paragraphs that follow describe an approach to the question of how to relate the measures just discussed to operational effectiveness. It is expected that this new approach will be the stepping stone to the problem of weighting the various weapon measures.

The firing engagements, whose results comprise the data base, are categorized according to the degrees of effectiveness as ~~defined by whether or not the firer achieved a hit during the quick-~~ fire engagement. Those engagements which required multiple trigger pulls (more than one round in semiautomatic or one burst of automatic) and failed to achieve a hit are defined as the least effective engagements; those which resulted in a hit whether single or multiple trigger pulls were required are defined as most effective engagements. A figurative presentation appears in Figure 4.

Level 3 is considered to be the center point on the scale since it contains all engagements observed during the quick-fire field experiment. Level 1 is the least effective set since it contains all engagements in which the firer pulled the trigger more than once and failed to achieve a hit. In other words, level 1 results when all multiple trigger pull engagements in which a hit did not occur are removed from level 3. Level 5 includes all engagements from level 3 that resulted in a hit on a target. Level 5 is defined as containing the most effective engagements and is the highest point on the effectiveness scale. Levels 2 and 4 are special cases comprised of engagements which can be placed intuitively on the scale somewhere between levels 1 and 3 and levels 3 and 5, respectively. Level 2 uses the same engagements for its data base as level 3, but all single trigger pull engagements have been removed. Since the targets disappeared when hit so that first round or burst hits terminated the engagement, the assumption is that soldiers who fired only once were apt to be more successful than soldiers who continued to fire. Level 4 consists of the same set of engagements that comprise level 5 except that one round or burst engagements that resulted in a hit have been removed. The assumption is that engagements that resulted in a hit but required repeated firings are less effective than single trigger pull engagements that resulted in a hit. The figure explains how the five levels are related. The scale is an ordinal scale which simply ranks the five levels. The distances between the levels on the scale are unknown, i.e., level 4 cannot be defined as four times more effective than level 1.

The number of measures which can be used to evaluate weapon performance varies from level to level. For example, the measure time-to-first-hit is only produced during engagements in which a hit occurred and, consequently, is not present in level 1. Figure 5 shows the measures that are associated with each of the five levels. Generally, as effectiveness increases more measures become available for the evaluation of weapon performance.

Once this effectiveness scale was developed, several measures were analyzed to determine how the measure related to combat effectiveness as defined by the effectiveness scale. An example that proved to be closely correlated to effectiveness was the size of an automatic burst of fire when a soldier engaged a target using the fully automatic mode of fire. Figure 6 shows burst size for the two test weapons plotted on the effectiveness scale. The curves show that for engagements with no hits there is relatively little difference as far as burst size is concerned. As engagements become more effective, Rifle A tends toward 2-round bursts while Rifle B tends toward the 3-round burst. This indicates that burst size plays an important role in the manner in which soldiers effectively employ rifles in the automatic mode. Soldiers who fire Rifle A in 2-round bursts get significantly more hits on targets while those who fire Rifle B in 3-round bursts achieve more hits. In statistical language, an interaction has occurred between real world effectiveness, as defined by the scale, and burst size.

This discovery, made through the detection of interaction effects (4), suggested that a more thorough analysis of burst size would lead to a better understanding of the factors which contribute to the combat effectiveness of the soldier/weapon system. The first step was to examine the size of the bursts which soldiers fired in the automatic mode. Figure 7 shows the results of each automatic mode trigger pull observed throughout the quick-fire experiment. The average burst size for Rifle A was 2.27 rounds; for Rifle B the average was 2.78 rounds. Burst hit probabilities for the two weapons are shown in Figure 8. The analysis thus far fails to show a difference in effectiveness between the two rifles. Consequently, further analysis of burst size was attempted.

The next effort was to examine the individual effectiveness of each round of the different burst sizes; these are shown in Figures 9 and 10. The highest single-round effectiveness for both weapons is achieved with the single round of a semiautomatic trigger pull. The second most effective round for Rifle A is the first round of a 2-round burst; the same phenomenon for the first round of automatic fire for Rifle B was not observed. It could be expected that all first rounds of a burst and the single round of a semiautomatic trigger pull would be identical in terms of the probability of achieving a hit. The effort required to acquire a target, position the weapon, align the sights, and squeeze the trigger are exactly the same regardless of the mode or the number of rounds that follow the first round. These data do not support this reasoning and, in fact, show that first-round effectiveness is related to the number of rounds that follow and the effectiveness of the following rounds. Generally, individual round effectiveness drops very rapidly as the burst of fire continues in a quick-fire engagement. Bursts of four or more rounds are relatively ineffective against point targets.

The comparative effectiveness of each burst or automatic trigger pull can be seen in Figure 11. This figure shows the cumulative hit probability of each round within the burst. Cumulative hit probabilities show very clearly that burst size is related to the ability to hit a point target. Without the burst analysis no difference between weapon systems was observed (see Figure 8). From the burst analysis, the most effective weapon system is Rifle B in the 3-round burst mode. Since firing techniques can be improved by training, the potential of Rifle B appears to be greater than that of Rifle A as an automatic fire weapon.

Two important findings have materialized from the burst analysis that were not among the original objectives of the quick-fire experiment: The Service Test must determine the weapon system's optimum operating mode to yield complete information on weapon potential, and training procedures are related directly to weapon performance and therefore should not be considered as separate entities. These are examples of the indirect benefits from a study based on a solid methodological foundation.

This technique of equating a specific measure to real world effectiveness represents an advance in military test procedures, but is limited in that the "real world" is still a simulated combat firing facility. Still, even with its limitations, the combat test facility is a dynamic test environment, a model that brings into play many of the influencing variables common to the combat environment, and, as such, provides a distinct improvement in operational testing. The extent to which these measures influence effectiveness can be used as a method of weighting the measures to permit proper emphasis when weapon-selection decisions must be made.

SUMMARY

All managers, whether in business, government, or the military, are faced with the same task: the accomplishment of their objectives with the most efficient utilization of resources. And, the problem that all managers have in common is to relate their decision criteria to their goals. For example, a corporation executive must relate such measures as cost of labor and material, labor turnover, training costs, and equipment depreciation to a thing called profit. A weapon designer or test officer must relate such measures as holes in targets, rate of fire, timeliness of fire, malfunctions, and cost of production to a thing called success in the combat environment. To accomplish this task managers must bridge the gap between what he can measure and what the system being measured is expected to accomplish in the real world. This paper describes an attempt to reduce this area of subjectivity by demonstrating a method which relates one measure to one scale of effectiveness. But, if one scale can be developed so can others. Perhaps scales can be developed which will relate time, malfunctions, and ammunition expenditure to weapon effectiveness on this dynamic testing model. Finally, as scales are developed they may eventually be related to each other. Then the measures which we refer to by the general categories of responsiveness, accuracy, reliability, sustainability can be equated against each other reducing to an absolute minimum the manager's need to make subjective decisions. If successful, the scaling technique, based on the detection of interactions, will provide the manager with a powerful new decision-making tool.

REFERENCES

(1) Klein, Ronald D., Final Report, Quickfire Experiment I, Infantry Weapons Test Methodology Study, U.S. Army Infantry Board, 27 June 1969.

(2) Jackson, John D., Small Arms Methodology Review, U.S. Army Infantry Board, December 1969.

(3) Klein, Ronald D. and Brown, Robert M., The Development of Combat Related Measures of Effectiveness for Small Arms Weapon Systems, U.S. Army Infantry Board, November 1969.

(4) Bargmann, Rolf E., Exploratory Techniques Involving Artificial Variables, A Presentation at the Second International Symposium on Multivariate Analysis, June 17, 1968.

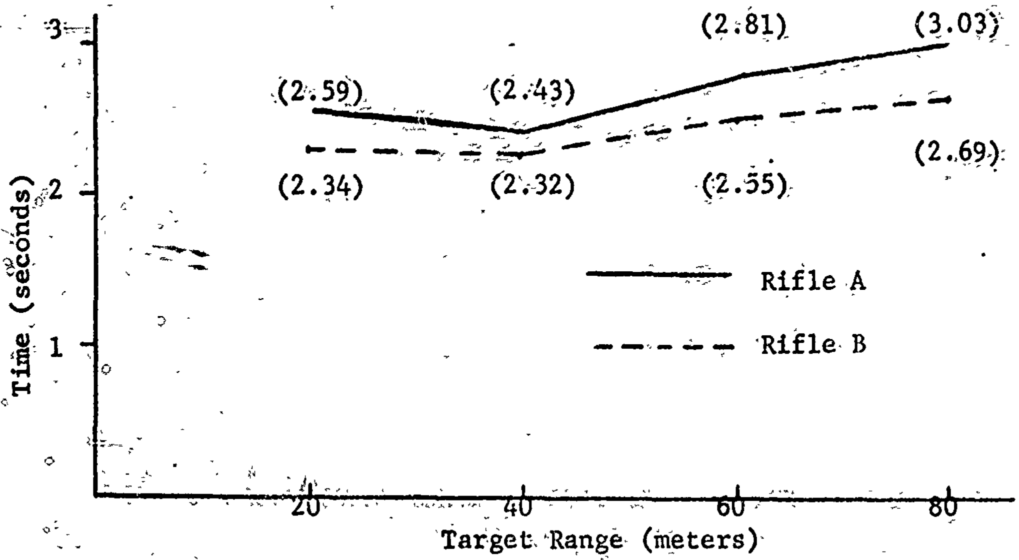


Figure 1

AVERAGE TIME TO FIRST ROUND (BOTH MODES)

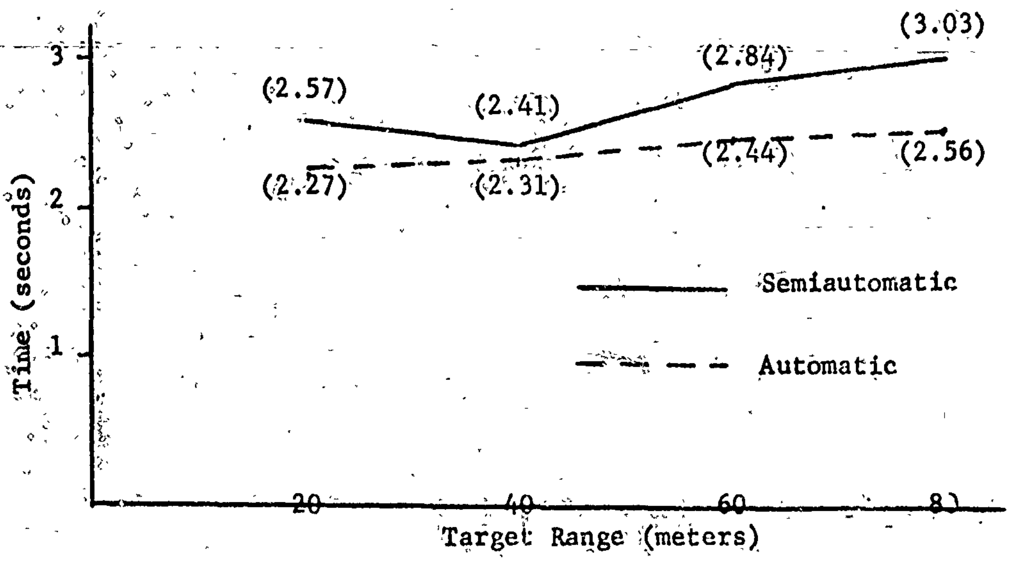


Figure 2

AVERAGE TIME TO FIRST ROUND (BOTH RIFLES)

	Rifle A	Rifle B	Significant?
All ranges	.76	1.04	Yes
20 meters	.66	1.00	Yes
40 meters	.72	.86	No
60 meters	.75	1.13	Yes
80 meters	.84	1.13	Yes

Figure 3

TIME BETWEEN BURSTS (SECONDS)

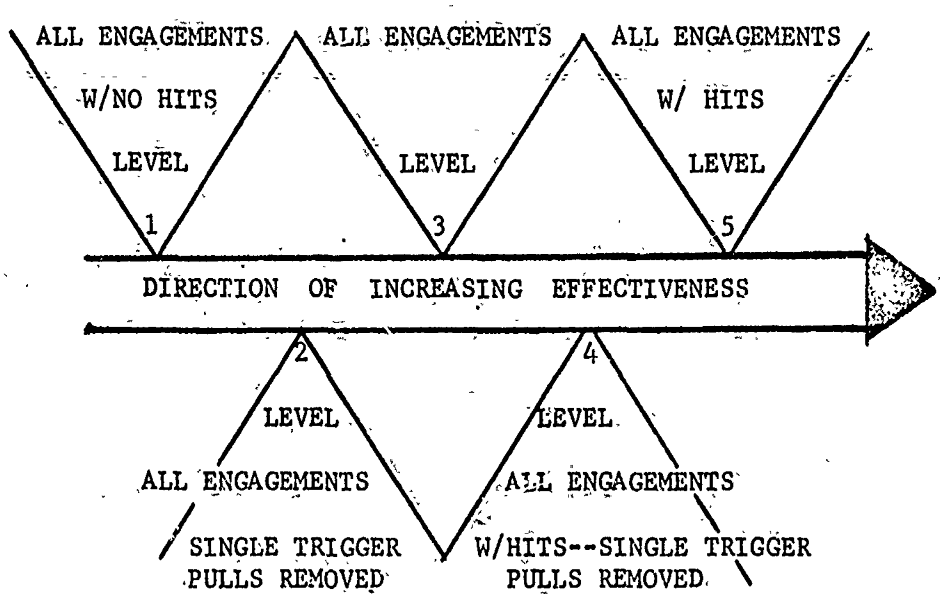


Figure 4

ORDINAL SCALE OF EFFECTIVENESS

INCREASING EFFECTIVENESS ↑

Level			
Level 5	Same as Level 4	Less	Time Between Trigger Pulls
Level 4	Same as Level 3	Plus	Time Between Trigger Pulls Time to First Hit Number of Rounds to First Hit
Level 3	Same as Level 2	Less	Time Between Trigger Pulls
Level 2	Same as Level 1	Plus	Single Shot Hit Probability Burst Hit Probability Number of Hits
Level 1	Time to First Round Number of Trigger Pulls Total Rounds Fired		Time Between Trigger Pulls Length of Burst--Automatic

Figure 5

MEASURES AVAILABLE AT EACH EFFECTIVENESS LEVEL

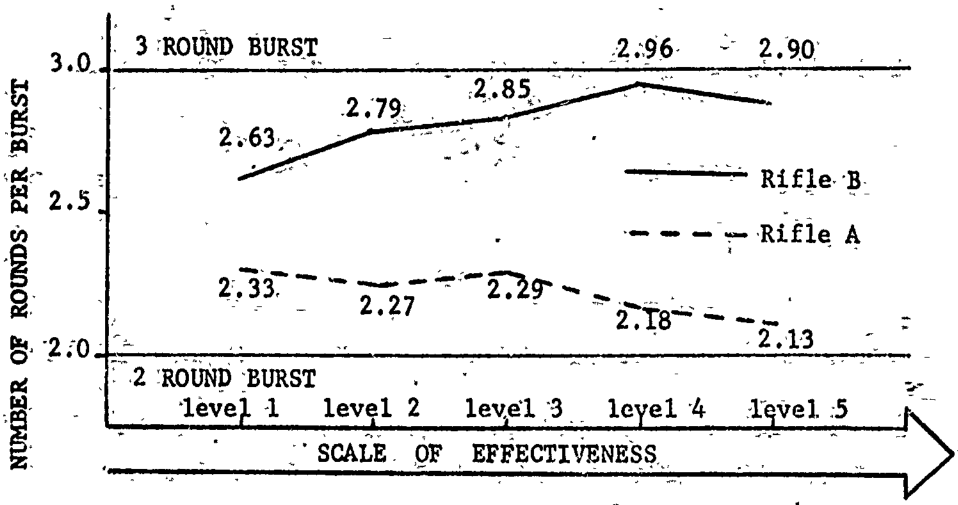


Figure 6

BURST SIZE VS. EFFECTIVENESS OF ENGAGEMENT

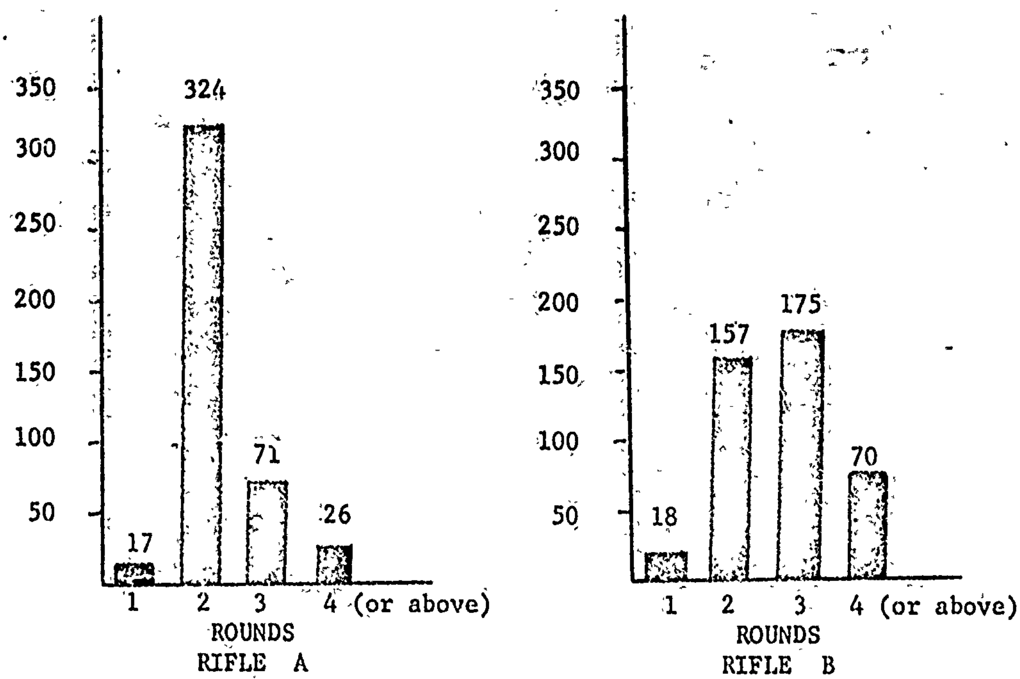


Figure 7

NUMBER OF OBSERVATIONS OF EACH BURST SIZE

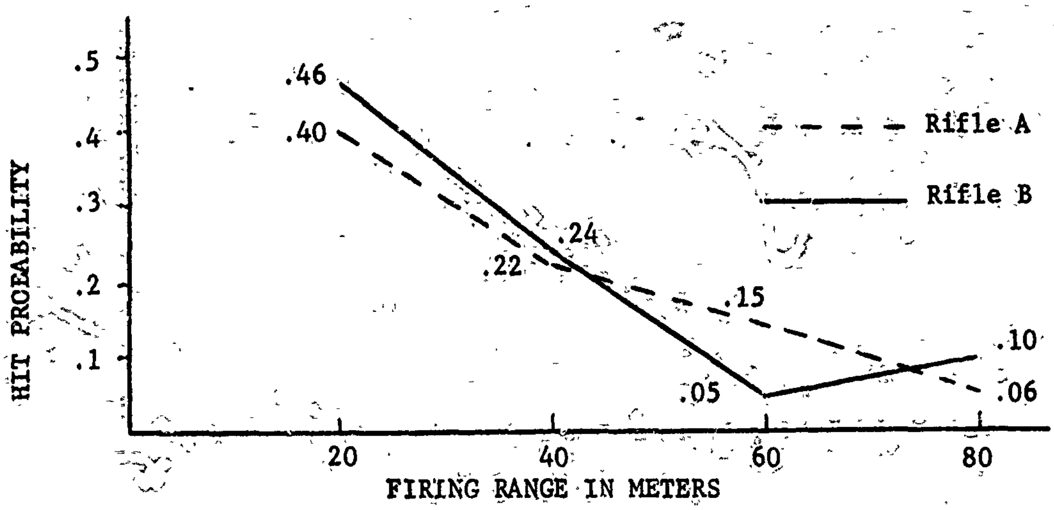
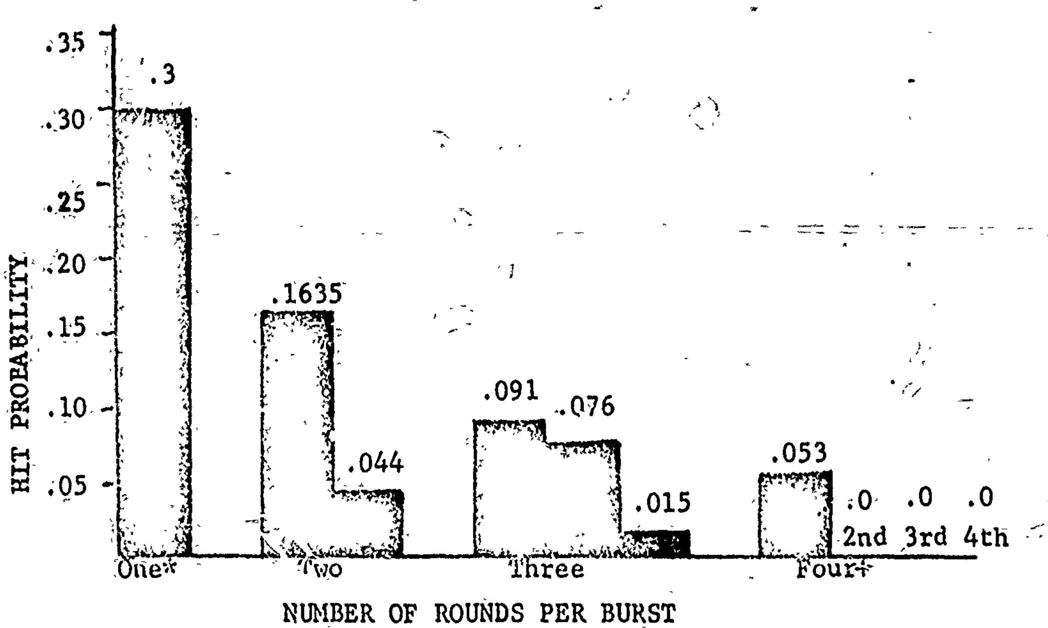


Figure 8

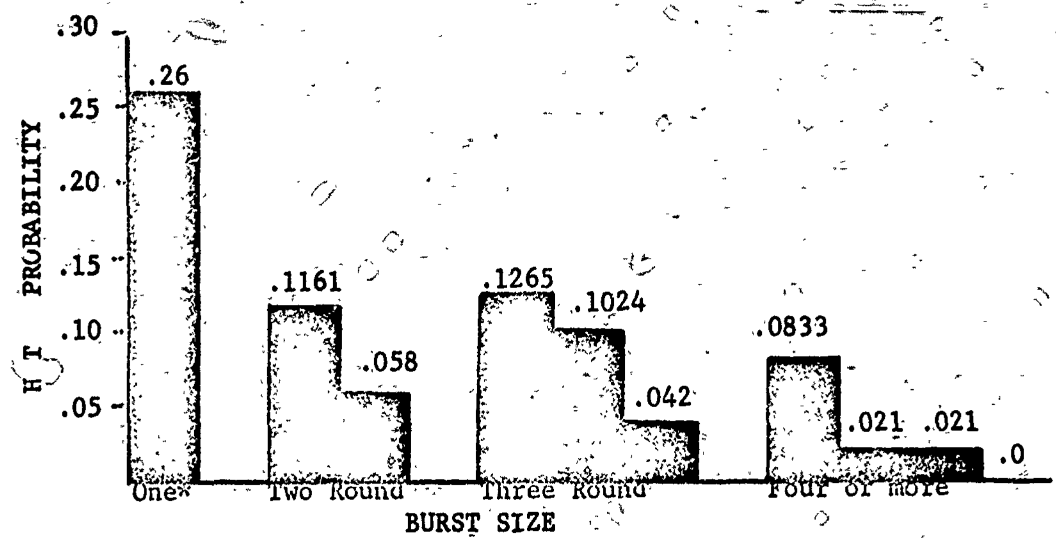
RIFLE COMPARISON -- BURST HIT PROBABILITY



*Includes all Semiautomatic Fire

Figure 9

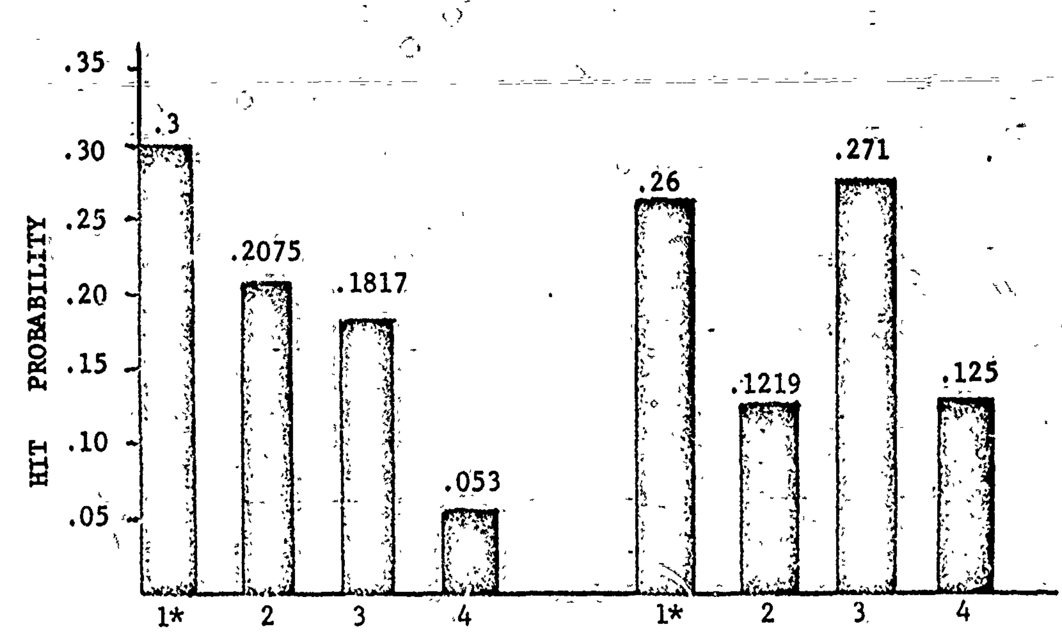
INDIVIDUAL ROUND HIT PROBABILITIES FOR RIFLE A



*Semiautomatic

Figure 10

INDIVIDUAL ROUND HIT PROBABILITIES FOR RIFLE B



*Semiautomatic ROUNDS PER BURST

RIFLE A Figure 11 RIFLE B

CUMULATIVE BURST HIT PROBABILITY